1990-03

An application of survival analysis methods to the study of Marine enlisted attrition

Hawes, Eric A.

Monterey, California: Naval Postgraduate School

http://hdl.handle.net/10945/34851

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

AN APPLICATION OF SURVIVAL ANALYSIS METHODS TO THE STUDY OF MARINE ENLISTED ATTRITION

by

Eric A. Hawes

March, 1990

Thesis Advisor: R. R. Read

Unclassified

security classification of this page

REPORT DOCUMENTATION PAGE

1a Report Security Classification Unclassified	1b Restrictive Markings
2a Security Classification Authority 2b Declassification/Downgrading Schedule	3 Distribution Availability of Report Approved for public release; distribution is unlimited.
4 Performing Organization Report Number(s)	5 Monitoring Organization Report Number(s)
6a Name of Performing Organization Naval Postgraduate School 6b Office Symbol (if applicable) 30	7a Name of Monitoring Organization Naval Postgraduate School
6c Address (city, state, and ZIP code) Monterey, CA 93943-5000	7b Address (city, state, and ZIP code) Monterey, CA 93943-5000
8a Name of Funding/Sponsoring Organization 8b Office Symbol (if applicable)	9 Procurement Instrument Identification Number
8c Address (city, state, and ZIP code)	10 Source of Funding Numbers Program Element No / Project No / Task No / Work Unit Accession No

11 Title (Include security classification) AN APPLICATION OF SURVIVAL ANALYSIS METHODS TO THE STUDY OF MARINE ENLISTED ATTRITION

12 Personal Author(s) Eric A. Hawes

13a Type of Report	13b Time Covered	14 Date of Report (year, month, day)	15 Page Count
Master's Thesis	From To	March 1990	90

16 Supplementary Notation The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 Cosati Codes			18 Subject Terms (continue on reverse if necessary and identify by block number)
Field	Group	Subgroup	military attrition, survival analysis, log-linear regression, theses, (CP)K

19 Abstract (continue on reverse if necessary and identify by block number)

This thesis is an application of survival analysis methods to study first term enlisted attrition from the Marine Corps. The data comprise over 99 percent of all enlisted accessions into the Marine Corps between 1 October 1983, and 31 August 1988. A large percentage of the observations are censored, thus motivating the use of survival analysis techniques. The enlistees are categorized by three covariates: education credential, Armed Forces Mental Group ~~(AFMG)~~, and presence/non-presence of a moral waiver. The attrition behavior of the enlistees is then examined to identify which covariate classifications are associated with premature attrition. The majority of the findings concerning the effects of the covariates on attrition are consistent with published results from previous military attrition studies. Two findings of the thesis, though, are perhaps new. First, the attrition behavior of alternate high school credential holders varied significantly according to credential type. Second, the relationship between aptitude and attrition behavior appears to have weakened in recent years.

The thesis also provides an opportunity to evaluate the uncommon practice of using survival analysis methods to examine military attrition. The results are promising as the survival analysis methods prove to be both accurate and efficient. Graphical plots of survivor function estimates provide an easily understood illustration of attrition behavior. The use of log-linear regression to model military attrition shows potential as a desk-top tool for conducting informal analyses.

Keywords:

20 Distribution Availability of Abstract ☒ unclassified/unlimited ☐ same as report ☐ DTIC users	21 Abstract Security Classification Unclassified
22a Name of Responsible Individual R.R. Read	22b Telephone (include Area code) (408) 646-2382 22c Office Symbol 55RE

DD FORM 1473,84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete

security classification of this page

Unclassified

An Application of Survival Analysis Methods to the Study of Marine Enlisted Attrition

by

Eric A. Hawes
Captain, United States Marine Corps
B.I.E., Georgia Institute of Technology, 1984

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1990

Author: [redacted]

Eric A. Hawes

Approved by: [redacted]

R.R. Read, Thesis Advisor

[redacted]

P.R. Milch, Second Reader

[redacted]

Peter Purdue, Chairman,
Department of Operations Analysis

ABSTRACT

This thesis is an application of survival analysis methods to study first term enlisted attrition from the Marine Corps. The data comprise over 99 percent of all enlisted accessions into the Marine Corps between 1 October 1983, and 31 August 1988. A large percentage of the observations are censored, thus motivating the use of survival analysis techniques. The enlistees are categorized by three covariates: education credential, Armed Forces Mental Group (AFMG), and presence/non-presence of a moral waiver. The attrition behavior of the enlistees is then examined to identify which covariate classifications are associated with premature attrition. The majority of the findings concerning the effects of the covariates on attrition are consistent with published results from previous military attrition studies. Two findings of the thesis, though, are perhaps new. First, the attrition behavior of alternate high school credential holders varied significantly according to credential type. Second, the relationship between aptitude and attrition behavior appears to have weakened in recent years.

The thesis also provides an opportunity to evaluate the uncommon practice of using survival analysis methods to examine military attrition. The results are promising as the survival analysis methods prove to be both accurate and efficient. Graphical plots of survivor function estimates provide an easily understood illustration of attrition behavior. The use of log-linear regression to model military attrition shows potential as a desk-top tool for conducting informal analyses.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

A. PROBLEM STATEMENT

Since the mid-1960's the Marine Corps has used education and aptitude data to screen individuals applying for service. Throughout the ensuing years, education level has proved to be far and away the best predictor of whether a recruit will complete the first term of enlistment. This is significant because the ability to complete the first-term of service is one of the yardsticks most often used by military planners to measure the success of recruits. Accordingly, the recruiting effort of all the services has centered on education and specifically the pursuit of high school graduates. The results have been positive. Premature attrition has been reduced and the overall quality of recruits has never been higher.

Many questions remain unanswered, however. For instance, it is not exactly clear *why* high school graduates are better suited for military service. Issues concerning the non-high school graduate population are even more troubling, due to this group's history of high attrition. Studies to date have fallen short in identifying background factors that indicate a better suitability for service among non-high school graduates.[1] Additionally, several programs through which an alternate high school degree may be attained have recently come into existence. Doubt currently resides within the Department of Defense that the recepients of these alternate degrees are as well suited for military service as those receiving a high school diploma in the traditional manner. Is this doubt well founded? If so, does it apply to all alternate high school credentials?

The Marine Corps has enjoyed, from a recruiting standpoint, its finest period ever during the 1980's. However, if the indicators are correct, rougher times lie ahead. With increasing frequency during the 1990's, the Marine Corps may be forced to recruit from the non-high school graduate population. While attrition rates will surely rise, the magnitude of the increase will be a function of the Marine Corps' ability to discriminate among non-high school graduates, and to identify those with the greatest probability of completing their initial term of service.

---

1 Some background factors that are typically examined are an individual's education level, aptitude, and presence of a moral waiver.

B. THESIS OVERVIEW

1. Purpose

This thesis examines the relationship between a recruit's covariates (education, aptitude, moral waiver status) and attrition behavior. Of course, one of the primary aims of the thesis is to identify those covariate values that are most prevalent among recruits that complete their first term of service. In this sense, the thesis is similar to the majority of military attrition studies of the last couple of decades. What is somewhat novel about this thesis is the use of survival analysis methods to approach the problem. Survival analysis techniques are used extensively in the biomedical sciences and are characterized by their ability to handle censored, or incomplete, observations. While conceptually, survival analysis methods enjoy certain advantages over traditional approaches to manpower studies, the practical application and assessment of their effectiveness and applicability to military attrition issues is limited at best.

2. Findings

For the most part, the findings of the thesis are consistent with conclusions previously reached in other military attrition studies. A couple of findings, though, seem to be unique and warrant further consideration. The first of these findings is that the attrition behavior of recruits with alternate high school credentials varies significantly according to credential type. The *Alternate High School Credentials* section of Chapter V addresses this issue. The second finding is that the usefullness of aptitude as a predictor of attrition behavior diminished somewhat in the more recent years. This issue is addressed in the *Single Covariate Effects* and *Post Boot Camp Attrition* sections of Chapter V.

The survival analysis techniques used proved to be both appropriate and effective. The primary tool used in the thesis is the *product-limit estimate* of the survivor function. In the presence of large sample sizes, the *product-limit estimate* provides a well-behaved, easy to interpret illustration of attrition behavior. Log-linear regression is used to model the attrition behavior. While the results yielded by the model are imperfect, the accuracy and efficiency are sufficient to justify the use of log-linear regression as an analytical tool for conducting informal attrition studies.

3. Organization

This thesis is arranged in the same order as the study was performed. Chapter II sets the stage for the thesis by providing a summary of previous attrition studies and current problems facing military recruiters. Chapter III establishes the definitions and methodology of the survival analysis techniques used in the thesis. Chapter IV defines

the covariates and contains an exploratory analysis of the data. Chapter V is the heart of the thesis, containing analyses of single and combined covariate effects as well as detailed analyses of the attrition behavior of alternate credential holders and boot camp graduates. Chapter VI contains the development, methodology, and evaluation of the log-linear regression model. Two examples of the model in use are provided. Chapter VII contains conclusions on the use of survival analysis methods to study military attrition, as well as a summary of the findings and recommendations for future study.

II. ATTRITION STUDIES

A. HISTORICAL BACKGROUND

1. Education Credential and Aptitude

Historically, non-high school graduates have been approximately twice as likely as high school graduates to prematurely attrite during their first term of service. Accordingly, a high school education has been the most highly sought quality in recruits by all of the U.S. Armed Forces. Initially, potential recruits could be easily classified into one of two categories; high school graduate or nongraduate. Non-high school graduates have been traditionally required to meet higher aptitude standards than their high school graduate counterparts. This follows from studies showing that even high school graduates with low aptitude scores are less likely to attrite prematurely than non-high school graduates with high aptitude scores. Therefore, aptitude has taken a backseat to education credential as a predictor of recruit attrition. The Armed Forces all use the Armed Services Vocational Aptitude Battery (ASVAB) to evaluate the academic and vocational abilities of applicants. The Armed Forces Qualification Test (AFQT) consists of four of the ten ASVAB subtests (Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Numerical Operations). AFQT scores are used to categorize individuals into an Armed Forces Mental Group (AFMG). AFQT scores by themselves have not been found to be a significant predictor of first-term attrition. However, when education level is held constant, there is a relationship between AFMG and attrition. Those in the higher aptitude groups demonstrate a lower attrition rate during the first term of service. However, the relationship is rather weak, particularly among non-high school graduates. Thus minimum aptitude standards have been established for different education groups not as a means of controlling attrition, but rather as a method of reducing the number of enlistment-eligible nongraduates. [Ref. 1]

2. Moral Waivers

A moral waiver is often required for enlistment if an individual has a history of criminal offenses or substance abuse. The relationship between moral waivers and attrition is not nearly as established as with education and aptitude. Recently, Fitz and McDaniel concluded that Marines with nontraffic moral waivers had a greater probability of being discharged for unsuitability than the rest of the population [Ref 2].

However, this relationship was not as strong for non-high school graduates, and studies on moral waivers thus far have produced as many questions as they have answered. The utility of using moral waiver criteria in the recruit selection process remains unclear.

B. MOTIVATION FOR THIS STUDY

1. Alternate Education Credentials

The recent advent of competency tests and alternate high school credentials have complicated the education classification system, as the lines for education categories are no longer clearly drawn. One of the earliest and perhaps most well known equivalency credential, the test-based General Education Development (GED), has become widespread enough to often be classified as a separate education category. GED holders, however, have behaved like non-high school graduates in terms of premature attrition [Ref. 3]. Thus the Marine Corps, for the purposes of recruit screening, have imposed more stringent aptitude requirements on GED holders than high school diploma graduates. Presently, the issue is more complex as both the numbers and types of alternate high school credentials are increasing.

2. DOD Education Codes and Time Spent in School

The Department of Defense education coding system classifies the holders of all recognized equivalency tests and alternate credentials into one category. This grouping is curious because considering the diverse chararcteristics of many of the credentials, it seems unlikely that the attrition behavior of this group would be homogenous. While some studies have shown that time spent in school is more pertinent to adjusting to military life than actually obtaining a diploma, the DOD codes make no allowance for time spent in school [Ref. 4]. An individual who attends twelve years of school but fails to graduate can often receive a certificate of high school attendance. The DOD codes group a high school certificate of attendance recipient together with an individual that obtains a GED after completing an unknown amount of school.

3. Decrease in the 17 to 21 Year Old Male Population

The 17 to 21 year old male population of the United States is projected to decrease until the mid-1990's. This obviously is cause for great concern among military manpower planners. In recent years, close to ninety percent of Marine Corps recruits have obtained a high school diploma in the traditional sense. With a smaller manpower pool from which to recruit, it is not an unreasonable assumption that this percentage will decrease.

4. Future Challenges

The Marine Corps, and the armed forces in general, face some stiff manpower challenges in the upcoming decade. A shrinking defense budget and manpower pool coupled with a healthy national economy places more importance than ever on recruiting individuals well suited for military service. Though not the subject of this study, other reports have analyzed the enormous cost to the military each year of accessing, training, and ultimately discharging individuals who fail to complete their initial term of service [Ref. 5].

Curiously, however, little is known about the largest group of offenders in terms of premature attrition, the non-high school graduate. The present attitude of the military towards the recruitment of nongraduates is understandably negative. Nongraduates are accepted as a last resort, and then priority is given to those displaying higher aptitude and/or some sort of high school equivalency. The preferential treatment is not granted because this group has proven itself more likely to be suitable for military service, but because the members of this group are simply more qualified to perform most jobs. In the absence of compelling criteria to separate the "low-risk" nongraduate from the "high-risk" nongraduate, this is the most rationale course of action. However, this is a rather defensive policy considering that over half of all nongraduates become successful service members. The determination of which factors make these nongraduates more suitable for military service could lead to a more effective selection policy. This problem has relevance, because concerning nongraduates, two things seem almost certain. First, based on the foregoing indicators, the number of nongraduates accessed into the Marine Corps during the 1990's will likely increase. Second, with increasing frequency, the non-high school graduates of the future will come equipped with some sort of alternate high school credential. In order to meet this challenge, the Marine Corps must attempt to understand the attrition behavior of nongraduates. Specifically, a determination needs to be made as to which qualities present in a nongraduate make that individual a more likely candidate to complete the first term of service.

III. METHODOLOGY

A. SURVIVAL ANALYSIS TECHNIQUES

Statistical methods for survival analysis have evolved largely from biomedical applications. These methods deal directly with failure time data and are characterized by their ability to handle unobserved failure times. The testing of new medical treatments often involves the application of the different treatments to patients, then the subsequent observation of these patients, and results in a set of failure time data. Typically a failure is characterized as death or reoccurrence of some condition. Of course, it would be impractical to wait until every observation has failed before making conclusions concerning the treatments. Thus, the experiments are normally terminated at some point in time and the failure times for those patients that have not "failed" are considered to be censored observations. The censored failure times provide useful information, but they must be handled differently than the actual failures. Survival analysis methods have been and still are being developed to exploit this type of data.

The case of recruit attrition is somewhat analogous to the biomedical experiment just described. The failure time of a recruit is simply the length of service prior to premature attrition. Recruits who successfully complete their enlistment obligation may be treated as censored observations.

1. Advantages

Survival analysis methods provide a highly accurate, and efficient means to examine recruit attrition. The ability to accommodate censored observations allows for greater analytical power and flexibility without compromising the ability to model the process. A more detailed description of some of the advantages to applying survival analysis methods to the study of recruit attrition are outlined below.

(1) Handles Censored Observations. Most analyses of recruit attrition use non-censored longitudinal data. Non-censored longitudinal data requires a complete service history of each recruit. Thus, all recruits in the data population must complete their terms of service prior to the analysis. While this is a tried and proven method it also produces results that are dated by at least four years. By terminating the observation period, say after two years, and treating the unobserved failure times as censored observations, analyses may be performed that are much more current.

(2) *Provides Statistical Inference Capability.* The methods used in this study utilize essentially the entire population. However, the methods can be applied to a sample of the population as well. Parametric and non-parametric modelling techniques have been developed to exploit censored failure time data. These methods are typically more computationally efficient than mainstream techniques such as probit and logit.

(3) *Allows for Early Separations.* Early separations and other special cases of recruit attrition might need to be categorized differently than conventional failures, lest they bias the results of the analysis. Treating these special cases as censored observations distinguishes them from the failures while still providing information to the analysis. This capability might become especially important if the recently forecasted military force cuts become a reality.

B. NON-EAS ATTRITION AND CENSORING

The term "non-EAS attrition" (EAS meaning end of active service) will be used interchangeably with "premature attrition" throughout this study. Non-EAS attrition is defined as any separation from the Marine Corps that is not a direct result of either the completion of service obligation (EAS attrition) or a Marine Corps early release program.

The length of time a Marine spends on active duty prior to non-EAS attrition will be referred to as a Marine's failure time. For the purposes of this study, the failure times of all Marines that are not victims of non-eas attrition are unobserved and thus considered to be censored observations. The censoring procedure may be likened to the testing of medical treatments on terminally ill patients. In the analysis of these treatments, the failures are those patients who expire prior to the completion of the test period. All other patients are considered successful and deemed to be censored observations because their ultimate survival times have not been determined explicitly. In applying the life-testing procedure to the problem of personnel attrition, the counterparts to these successful patients, and thus the censored observations, are Marines that fall into one of two categories:

- those still on active duty at the end of the observation period.
- those that successfully completed their initial enlistment (EAS attrition).

Note that the second group above includes those Marines that obtained early releases from their enlistment contracts. Since early release programs are by definition

amicable agreements between the government and the Marine, these premature separations are not characterized as failures in this study.

C. COMPUTATION OF SURVIVOR FUNCTIONS

In this study it will be assumed that the failure times of Marines are discrete and represented by, t, $(t = 1,2,\ldots,51)$, where t is the number of months on active duty prior to failure (non-EAS attrition). The values of t are rounded to the next highest month. Therefore, if a Marine is discharged for an adverse reason after only nine days of active duty, the failure time for that Marine is one month. In the case where there are no censored observations, the computation of the survivor function, $S(t)$, at some time t is direct:

$$S(t) = \Pr\{\text{survival beyond time } t\},$$

$$S(t) = 1 - \left(\frac{\text{number of observations with failure time} \leq t}{\text{total number of observations}} \right). \quad (1)$$

The presence of censored observations, however, prohibits computing the exact survivor function. All that can be safely assumed about an observation censored at time t' is that the unobserved failure time is greater than t'. Thus equation (1) above cannot be solved when there are observations censored at values less than t. This problem is resolved by obtaining an alternate formulation of the survivor function using hazard rates. In the discrete case the hazard rate at some time t is defined as the conditional probability of failure at t:

$$\lambda_t = \Pr\{\text{an observation fails at time } t \mid \text{the observation survives to time } t\},$$

$$= \frac{\text{number of observations that fail during } (t-1,\ t]}{\text{number of observations with failure time} > (t-1)}. \quad (2)$$

While it might appear at first that the denominator of equation (2) violates the logic of the conditioning statement, this is not the case. For an observation to survive to at least a time t, that observation only has to survive an observable increment of time, a day in this case, past time $t-1$. This is due to the forementioned practice of rounding

the failure times to the next highest month. Using hazard rates, the survivor function may be expressed as:

$$S(t) = \prod_{j=1}^{t} (1 - \lambda_j). \tag{3}$$

An estimation, $\hat{\lambda}_j$, of the λ_j's would allow for an estimation of the survivor function, $\hat{S}(t)$, using equation (3). Let d_j be the number of failures at time j, $(j = 1, \ldots, t)$, and m_j be the number of observations censored in the interval $[j, j+1)$ for $(j = 0, \ldots, t)$. The number of observations at risk at a time just prior to j would be $n_j = (m_j + d_j) + \ldots + (m_t + d_t)$. It follows that $\hat{\lambda}_j = \frac{d_j}{n_j}$, $(j = 1, \ldots, t)$, and the *product-limit estimate* of the survivor function is:

$$\hat{S}(t) = \prod_{j=1}^{t} \left(\frac{n_j - d_j}{n_j} \right). \tag{4}$$

This estimate was derived by Kaplan and Meier (1958), and in practice is often referred to as the *Kaplan-Meier estimate* [Ref. 6]. It can be shown that (4) provides a nonparametric maximum likelihood estimate of the survivor function [Ref. 7: pp. 12-14]. The *product-limit estimate* is a highly developed and utilized method. Although not required in this study, methods have been developed to approximate confidence bands for (4). One such technique is presented by Gillespie and Fisher [Ref. 8].

IV. DATA OVERVIEW

This chapter provides a description of the data population and of each covariate and the possible covariate values. Additionally, the distributions of covariates are examined among the different cohorts; a cohort being defined as the population of recruits that enter the Marine Corps during a given fiscal year. This exercise provides an evaluation of the completeness of the data set as well as identifying trends and dissimilarities among cohorts. Lastly, a study of homogeneity is performed to help determine the best data-grouping strategy for the survival analysis.

A. POPULATION

This study involves male first-term enlistees with no prior military service (MFT-NPS) that were accessed into the the Marine Corps between 1 October 1983 and 31 August 1988. The MFT-NPS population accounts for approximately 90 percent of the total Marine Corps enlisted accessions during this period. As seen in Table 1 on page 12, the number of observations for each cohort is around 30,000. The exception of course is 1988, which is curtailed and thus has only about half the number of observations. The data used in this study encompesses over 99 percent of all MFT-NPS accessions into the Marine Corps during the forementioned period.

B. COVARIATES

1. Mental Group

Marine recruits are assigned an Armed Forces Mental Group (AFMG) based on their score on the AFQT. Appendix A contains the scoring requirements for each mental group. There are seven mental groups (I, II, IIIA, IIIB, IVA, IVB, V), and essentially the higher the score on the AFQT the lower the corresponding mental group number. Thus, AFMG I contains the highest aptitude personnel and AFMG V the lowest.

Table 1 on page 12 shows that the distribution of recruits among mental groupings was fairly even among the cohorts. Note that the top entry in each cell of the table is the total count and the number below , in parenthesis, is the cohort (row) percentage. This convention will be used throughout this report. The only discernable pattern evident from Table 1 is that the numbers of recruits in AFMG's I and II seems to be rising gradually during the period while the numbers in AFMG's IVA, IVB, and

V seem to be on the decline. This general increase in the aptitude of the recruit population is a predictable result of a increasingly better educated recruit population as discussed in the next section. There were 62 AFMG "unknowns" interspersed throughout the five cohorts. Since the number is so small and their placement appears to be random, their exclusion from the mental group survival analysis should not bias the results.

Table 1. RECRUIT DISTRIBUTION BY COHORT AMONG AFMG

	I	II	IIIA	IIIB	IVA	IVB	V	Total
84	863 (2.39)	11635 (32.25)	8440 (23.39)	13633 (37.79)	1446 (4.01)	24 (0.07)	21 (0.06)	36079
85	825 (2.85)	8172 (28.23)	6207 (21.44)	12685 (43.83)	1003 (3.47)	18 (0.06)	8 (0.03)	28944
86	849 (2.80)	9927 (32.71)	8225 (27.10)	11257 (37.10)	77 (0.25)	3 (0.01)	0 (0.00)	30346
87	1019 (3.22)	10885 (34.42)	8893 (28.12)	10733 (33.94)	87 (0.28)	2 (0.01)	0 (0.00)	31626
88	473 (3.30)	5011 (35.00)	3753 (26.21)	5047 (35.25)	28 (0.20)	3 (0.02)	0 (0.00)	14319
Total	4029 (2.85)	45630 (32.29)	35518 (25.13)	53355 (37.76)	2641 (1.87)	50 (0.04)	29 (0.02)	141314

2. Education Credential

Recruits are classified into one of three categories according to the Department of Defense Educational Codes implemented in October of 1987:

- TIER 1 - High school diploma graduate
- TIER 2 - Recognized alternate high school credential
- TIER 3 - Non-high school graduate

Appendix A contains more detailed documentation concerning the alternate credentials recognized under the TIER 2 classification.

The distribution of recruits by DOD educational code is provided in Table 2 on page 13 for each cohort. This time, however, the distribution between cohorts is not as even. For cohorts 86 to 88, the large number of "unknowns" in relation to the numbers of TIER 2 and TIER 3 recruits brings the usefullness of the data into question. The education data for cohort 88, with 569 education "unknowns" will not be utilized in this study. The reliability of the education data from cohorts 86 and 87 is largely

dependent on the attrition behavior of the "unknown" population. If as a group the "unknowns" demonstrate attrition patterns similar to the rest of the cohort as a whole, it can be assumed that the "unknowns" represent a random sample and thus their exclusion from the study will not bias the results. An examination of the randomness of the "unknowns" and subsequent discussion is contained later in the *single covariate effects* of this chapter.

Another notable aspect of Table 2 is that the percentage of TIER 1 recruits, or high school diploma holders, saw a steady rise during the observation period. While the presence of the "unknowns" somewhat clouds the issue, it also appears that the percentage of TIER 2 recruits is increasing relative to TIER 3 recruits during the period.

Table 2. RECRUIT DISTRIBUTION BY COHORT AMONG DOD EDUCATION CODES

	TIER 1	TIER 2	TIER 3	Unknown	Total
84	31004 (85.93)	3634 (10.07)	1431 (3.97)	10 (0.03)	36079
85	25464 (87.98)	2555 (8.83)	888 (3.07)	37 (0.13)	28944
86	27038 (89.10)	2450 (8.07)	659 (2.17)	199 (0.66)	30346
87	28304 (89.50)	2581 (8.16)	555 (1.75)	186 (0.59)	31626
88	13068 (91.26)	600 (4.19)	82 (0.57)	569 (3.97)	14319
Total	124878 (88.37)	11820 (8.36)	3615 (2.56)	1001 (0.71)	141314

3. Moral Waivers

The covariate of moral waiver is binary in nature; a recruit either does or does not possess a moral waiver. While a complete listing of the types of moral waivers granted is provided in Appendix A, no allowance is made as to moral waiver type in this study. Moral waivers are typically granted at one of three levels depending on the circumstances surrounding the event(s) that necessitate the granting of the moral waiver in the first place. The three levels are the recruiting station (RS), the recruit depot, and the Commandant of the Marine Corps (CMC). The majority of moral waivers granted by the Marine Corps involve rather minor offenses and are thus authorized at the lowest,

or RS level. This study will not examine these RS-granted moral waivers. Only those moral waivers granted by the recruit depot or CMC will be considered. Essentially, this places the emphasis on those recruits that received moral waivers for the offenses that are considered to be more serious in nature.

Table 3 below shows the distribution, by cohort, of moral waiver personnel. Note that for cohorts 86, 87, and 88 the percentage of recruits granted a moral waiver is around 7 to 11 percent. However, only two waivers appear in 84 and only 4.44 percent of the recruits in 85 are shown to possess moral waivers. Obviously the moral waiver data for 84 is incomplete and will therefore be excluded from this study.

Table 3. DISTRIBUTION OF MORAL WAIVERS BY COHORT

	Moral Waiver	No Moral Waiver	Total
84	2 (0.01)	36077 (99.99)	36079
85	1286 (4.44)	27658 (95.56)	28944
86	2224 (7.33)	28122 (92.67)	30346
87	3075 (9.72)	28551 (90.28)	31626
88	1500 (10.48)	12819 (89.52)	14319
Total	8087 (5.72)	133227 (94.28)	141314

One notable trend present in the moral waiver data is that the numbers of moral waiver personnel seems to be increasing over time. This may prove to be the case, at least concerning moral waivers granted at the depot and CMC levels, however given the amount of noise in the moral waiver data it's difficult to make any solid conclusions based on this set of observations.

The completeness of the moral waiver data for cohort 85 is suspect as well. Since the moral waiver data is almost nonexistent prior to cohort 85 and appears to be substantive after 85, the data was examined month by month for the entire 85 cohort to determine at what point in time, if any, the moral waiver percentages become consistent

with the remaining (later) cohorts. Table 4 on page 15 gives the distribution of moral waivers, by month, for cohort 85.

Table 4. FREQUENCY OF MORAL WAIVERS BY MONTH FOR COHORT 85

	Moral Waiver	No Moral Waiver	Total
Oct 84	1 (0.03)	3096 (99.97)	3097
Nov 84	2 (0.08)	2645 (99.92)	2647
Dec 84	1 (0.05)	2025 (99.95)	2026
Jan 85	17 (0.67)	2509 (99.33)	2526
Feb 85	130 (6.02)	2028 (93.98)	2158
Mar 85	138 (7.17)	1788 (91.28)	1926
Apr 85	165 (8.72)	1728 (91.28)	1893
May 85	81 (6.41)	1182 (93.59)	1263
Jun 85	174 (6.53)	2490 (93.47)	2664
Jul 85	168 (5.92)	2672 (94.08)	2840
Aug 85	195 (6.51)	2800 (93.49)	2995
Sep 85	214 (7.36)	2695 (92.64)	2909
Total	1286 (4.44)	27658 (95.56)	28944

As Table 4 demonstrates, February 1985 sees a sharp increase in the number of moral waiver personnel. In fact, from February onward, the percentage of moral waiver personnel is 6.78 percent, reasonably close to the 7.33 percent moral waiver personnel for cohort 86. Therefore, in examining moral waivers in this study, data for recruits that began active duty prior to February 1985 will be excluded.

C. HOMOGENEITY

A major issue to be resolved prior to performing the analysis was just how to group the data. The data could be grouped across cohorts and treated as one large sample, or the cohorts could be examined individually. If the data was to be grouped into one sample, homogeneity between the cohort groups had to be established. A homogeneous group is a collection of individuals who are similar with respect to all known factors which affect propensity to leave [Ref. 9: p. 13]. The period of study (Oct 83 to Aug 88) did not see any obvious policy changes or other factors (e.g., large-scale war) that might affect certain cohort groups more than others. Therefore, homogeneity between cohort

groups is entirely plausible if it can be shown that the composition of the cohort groups, in terms of the forementioned covariates, are essentially the same and that the survival distributions are similar between all cohorts.

As demonstrated by the frequency tables on the previous pages, the composition of the cohorts in terms of the covariate values are not identical. More importantly, though, is that the dissimilarities between cohorts are not major. It is a reasonable assumption that the quality of recruit, in terms of the covariates, is not radically different from year to year. Of course, while the incomplete moral waiver data for cohorts 84 and 85 warrants some careful handling of the data, it is not an indicator of dissimilarities between the members of different cohorts.

Having established sufficient similarity with respect to the composition of the cohorts, the next step was to compare the distributions of the estimated survivor functions. Figure 1 on page 17 shows the estimated survivor functions by cohort of all MFT-NPS recruits. The greatest disparity between the cohorts occurs during the first four months of service. After the four-month point the distributions appear quite similar. Clearly, the cause of the large variability during the first few months of service is recruit failure during boot camp. There are many possible explanations for these differences in boot camp attrition from year to year despite a rather consistent quality of recruit throughout the period. Manning requirements coupled with the volume of recruits are arguably the major source of the variability. The vagrancies of boot camp attrition, however, are not the subject of this study. The resulting disparities in the attrition behavior of the cohorts must be addressed, though.

Due in part to the differences in attrition behavior, the cohorts will be examined, at least initially, separately in this study. Other compelling reasons exist for pursuing this approach. The established incompleteness of the moral waiver data and a potentially similar problem involving some of the education data will require some partitioning of the data independent of homogeneity considerations. Also, examining the cohorts individually allows for a more powerful analysis. Behavioral trends can be tested for among different cohorts vice one sample. This will help clarify if attrition behavior is remaining constant or changing in some manner with time.

At times during the study, it may be advantageous to pool the data over different cohorts. This will be done only after an exhaustive analysis has been performed on the individual cohorts and covariates involved to insure that the attrition behavior is reasonably similar among the cohorts to be pooled.

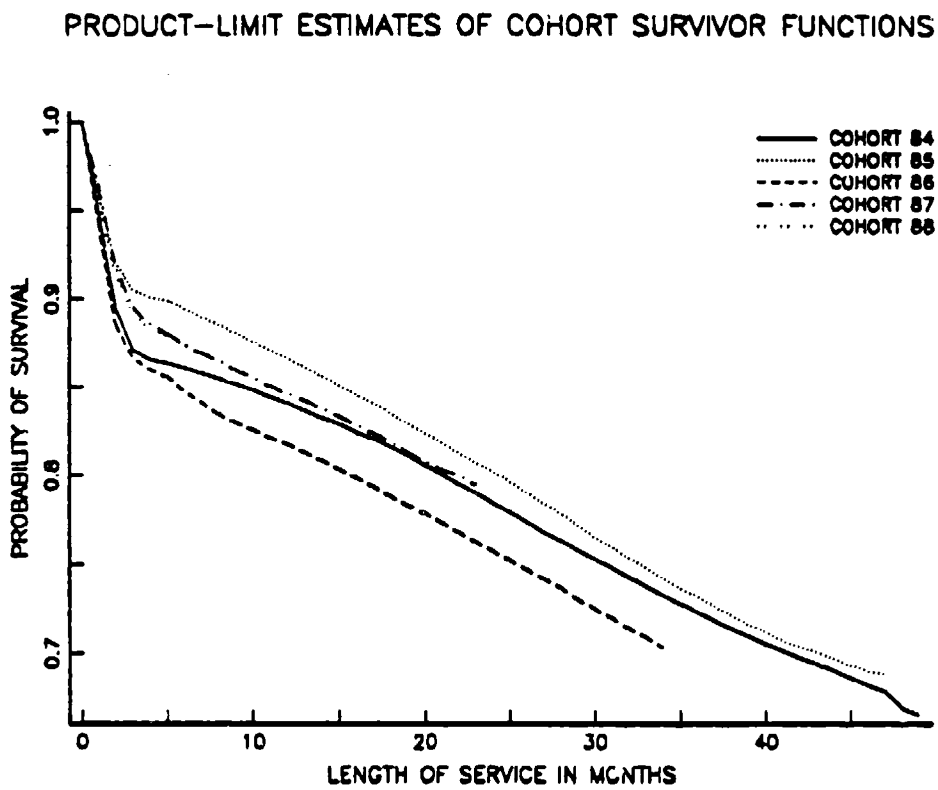

Figure 1. Comparision of Cohort Survivor Functions

The survivor functions shown in Figure 1, and throughout the study, were computed using the SAS (Statistical Analysis System) procedure LIFETEST [Ref. 10: pp. 529-557]. The output generated by the SAS software for the survivor functions in Figure 1 is included as Appendix C.

V. SURVIVAL ANALYSIS

This section deals with the comparision of survivor functions after the cohort populations have been partitioned according to covariate values. The individual effects of each covariate are examined at first, then the covariates are grouped and studied for joint effects. The final two sections of the chapter examine the attrition behavior of alternate high school credential holders and boot camp graduates. As the data are so complete, the comparative analyses are possible by graphical study of the survivor functions. In many instances, the attrition behavior is almost identical among cohorts. For the sake of brevity, when this occurs Cohort 84 survivor functions will illustrate the behavior in the body of the report and the remaining cohorts' survivor functions will be included in Appendix B.

Sample sizes are not indicated in this study. In all instances where an estimate of the survivor function is presented, the sample sizes are sufficiently large that statistical inference is not an issue. In other words, if two survivor functions appear to be different when plotted side by side, then statistically speaking, they are different at any reasonable significance level. Some sample sizes were too small to meet this criteria. In these instances either a qualifying statement is made in the text or the survivor function is omitted.

A. SINGLE COVARIATE EFFECTS

1. Education

Consistent with past studies of recruit attrition, education level proved to be the most significant indicator of a recruit's suitability for military service. Figure 2 on page 20 shows the estimated survivor functions, by DOD education code, for cohorts 84 and 85. Figure 3 on page 21 contains the same for cohorts 86 and 87. Cohort 88 is omitted, as discussed previously, due to the large number (almost 4 percent) of education "unknowns" present in that cohort.

From Figures 2 and 3 it is clear that TIER 1 recruits (HS diploma graduates) demonstrate the greatest survivability and TIER 3 recruits (non-HS graduates) represent the highest attrition risk group. The only notable dissimiliarity between cohorts is the attrition behavior of the TIER 2 recruits (alternate HS credentials). While for all cohorts the survivor function for TIER 2 recruits lies between the TIER 1 and TIER 3 survivor functions, its placement varies somewhat. As an example, consider cohorts 85

and 87. For cohort 85, the attrition behavior of the TIER 2 recruits more closely resembles the behavior of the TIER 1 recruits than the TIER 3 recruits. However, For cohort 87, the TIER 2 recruits more closely resemble the TIER 3 recruits in terms of attrition.

This variability warrants a closer look at the alternate high school credentials. The TIER 2 population is composed of recruits with one of several alternate HS credentials. Since these credentials are diverse, the background and social experiences of the TIER 2 recruits are presumably diverse as well. A significant degree of heterogeneity within the TIER 2 population could explain the variability of the survivor function. If certain credential holders exhibit greater survivability than others, a change in the composition of the TIER 2 population could cause large changes in the TIER 2 survivor function. The most effective means to study this possibility is to subdivide the TIER 2 population according to their actual credentials. This is performed later in the *Alternate High School Credentials* section of this chapter.

Figure 24 on page 61 of Appendix B is a duplication of Figure 3 with the survivor functions for the education "unknowns" included. The "unknown" population displays an extremely high attrition rate. This behavior suggests that the education "unknown" population is not a result of random data gaps. It appears that the education data of recruits who attrite after a short period of service is less than complete. Since TIER 2 and TIER 3 recruits have an established pattern of higher attrition than TIER 1 recruits, it stands to reason that the "unknown" population contains a disproportionately large number of non-high school diploma graduates. This of course will bias the survivor functions for cohorts 86 and 87. In the case of Figure 3, the bias is probably insignificant for TIER 1 and TIER 2 recruits because the number of "unknowns" is small compared to the subgrouping sizes. However, this is not the case for TIER 3 recruits. The number of "unknowns" is about one-third the number of TIER 3 recruits in cohorts 86 and 87. This will likely inflate the survivor functions of the TIER 3 recruits to a degree. While an inspection of Figure 3 doesn't suggest any problems, care should be taken when making empirical comparisions between cohort survivor functions.

2. Mental Group

High aptitude recruits have proved to be more likely to complete their first term of service. For each cohort the armed forces mental group number was inversely correlated with survivability. Figure 4 on page 22 represents cohort 84 and demonstrates this relationship. Note that AFMG I recruits are represented by the top curve, AFMG II

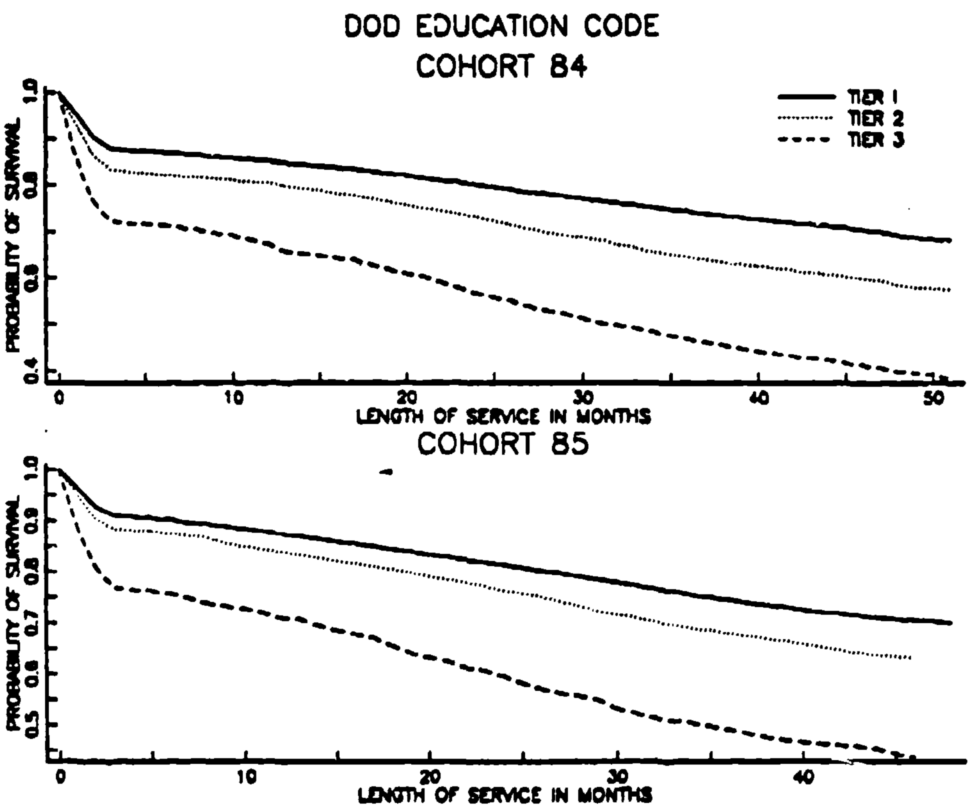

Figure 2. Survivor Functions by DOD Education Code, Cohorts 84 and 85

recruits by the next highest curve and so on to the lowest curve which represents the aggregate of AFMG IVA, IVB, and V recruits.

Based on these results it is obvious that aptitude on its own has some value as a predictor of recruit suitability. Of course in the case of the higher aptitude recruits (AFMG I and II), it is probable that the overwhelming majority are also high school diploma graduates. A more meaningful evaluation of the usefullness of aptitude as a

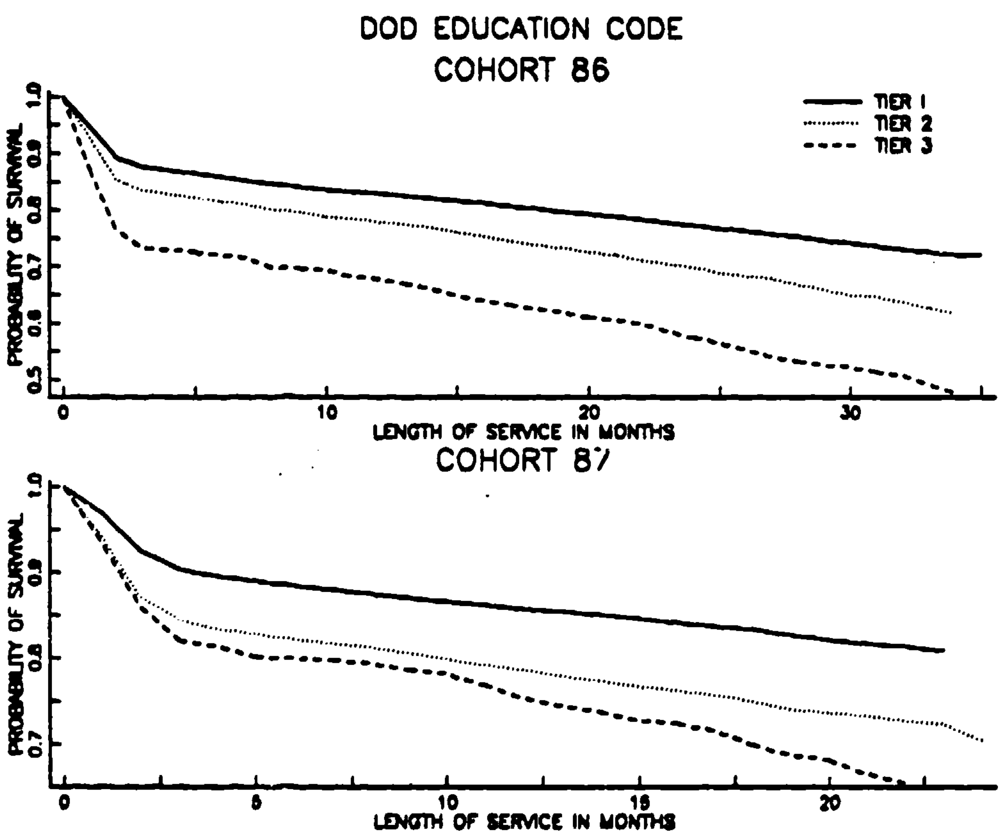

Figure 3. Survivor Functions by DOD Education Code, Cohorts 86 and 87

predictor is provided in the next section of this report when education and aptitude are considered jointly. At this point it is unclear whether a recruit's mental group is providing any information different from what is already provided by the DOD education code.

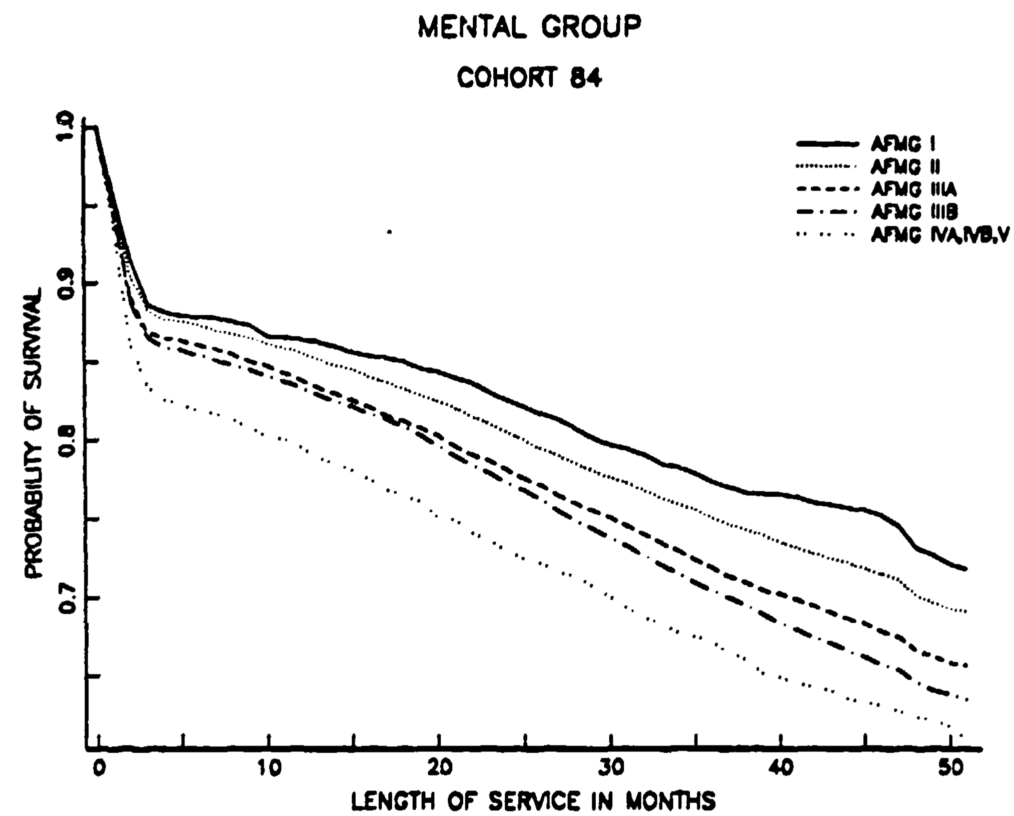

Figure 4. Survivor Functions by AFMG for Cohort 84

The estimates of the survivor functions, by mental group, for cohorts 85 through 88 are displayed in Figure 25 to Figure 28 of Appendix B. Note that the relationship between aptitude and attrition shown in Figure 4 for cohort 84 holds for these cohorts as well. However, it appears that the survivor functions of the different mental groups are grouped closer together in the later years. This potential trend is addressed in greater detail later in this chapter.

3. Moral Waivers

Recruits possessing a moral waiver had a slightly higher probability of attriting prematurely than those without a moral waiver. Figure 5 on page 24 and Figure 6 on page 25 contain the estimated survivor functions for cohorts 85 through 88. Note that the attrition behavior of the cohort 85 moral waiver personnel is different than that displayed within the other cohorts. This inconsistent behavior is likely related to the forementioned problem of missing moral waiver data for cohort 85 (recall that cohort 85 moral waiver data for recruits entering service prior to February 1985 was excluded from this study). Note in Figure 5 that the attrition for cohort 85 moral waiver personnel is inordinately low during the first four months of service. After the first four months, however, the behavior, in terms of the slope and thus hazard rate, is not unlike the moral waiver survivor functions for the other cohorts. It appears, much like the case with the education "unknowns", that moral waiver data has been omitted on a portion of the cohort 85 recruit-failure population. Therefore, the cohort 85 data will be excluded from further study involving moral waivers.

It is difficult to identify any notable trends concerning the attrition behavior of moral waiver personnel. While the presence of a moral waiver is clearly an indicator of higher attrition, the magnitude of this difference is small when compared to the differences in attrition among the various education and mental groups.

B. COMBINED COVARIATE EFFECTS

Since education is the dominant predictor of attrition among the three covariates, it is of interest to investigate the affects of the other covariates with education held constant.

1. Education by Mental Group

In the sections to follow, the populations for cohorts 84 through 87 are partitioned according to DOD education code. The three education subgroupings are then further broken down into armed forces mental group. This allows a couple of important issues to be addressed:

- Are the effects of aptitude consistent among education groups?
- Are the effects independent of cohort, or are new behavioral trends developing?

The second question can be examined since each cohort is considered separately. The issue of trends may be particulary important among the TIER 2 population since this group appears to be the most unstable in terms of attrition behavior.

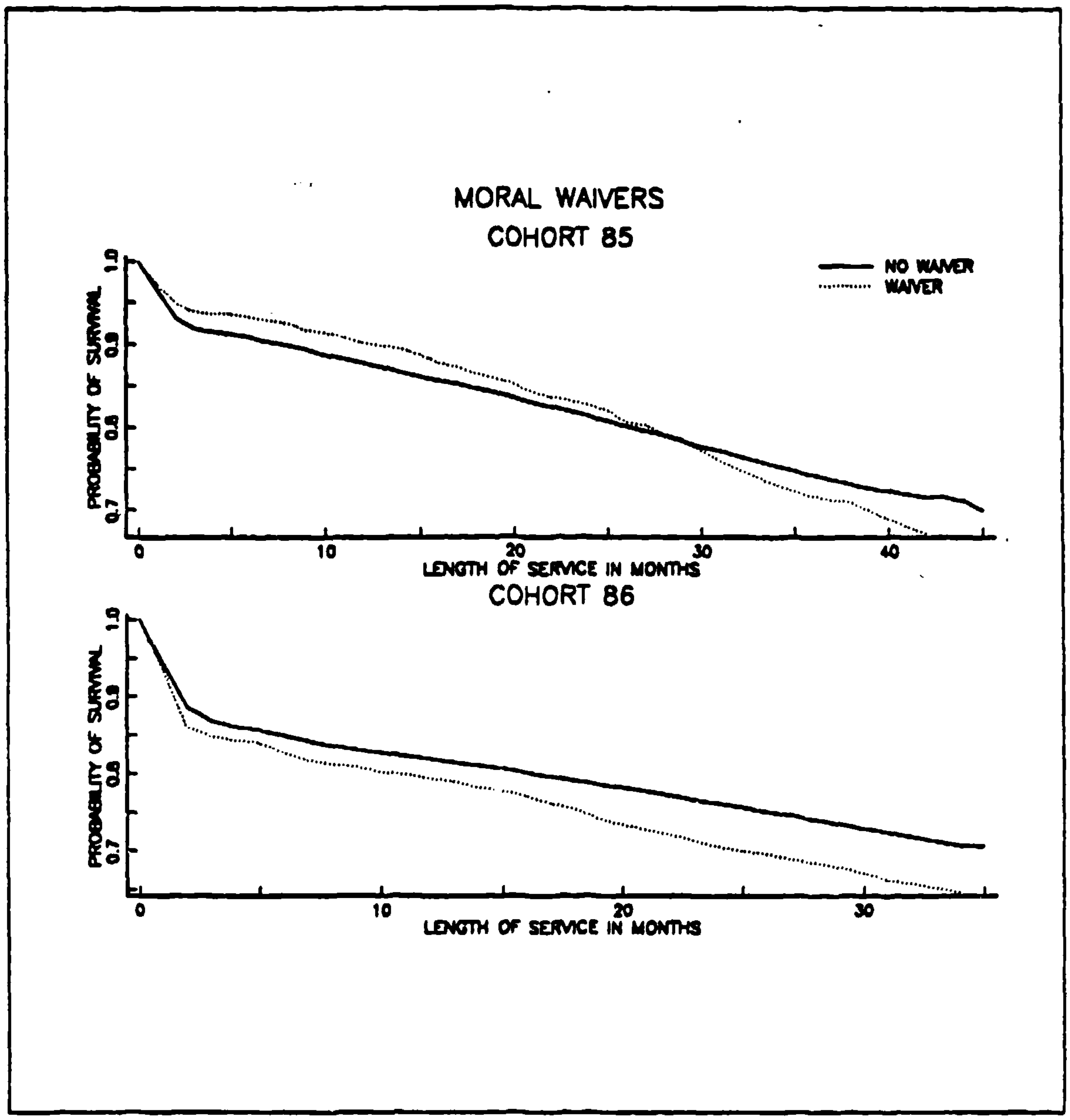

Figure. 5. Survivor Functions by Moral Waiver, Cohorts 85 and 86

The final section in the *Education by Mental Group* analysis involves the grouping of the four cohorts into one population. The population is then partitioned as before, into all possible combinations of education and mental group and a survival analysis is performed on each subgrouping. The result is a tabular comparison which ranks each subgroup according to survivability, from highest to lowest.

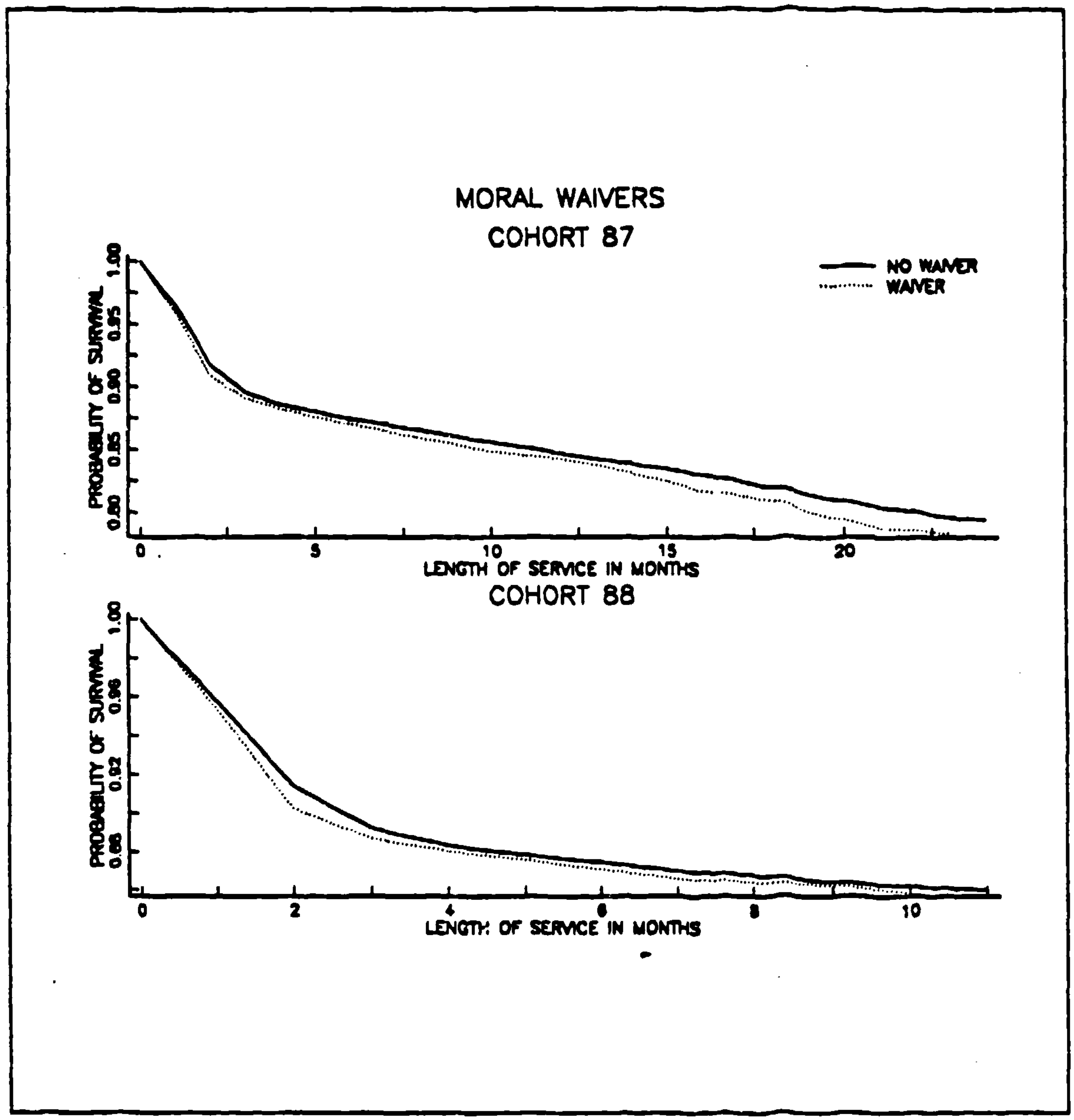

Figure 6. Survivor Functions by Moral Waiver, Cohorts 87 and 88

a. TIER 1

Figure 7 on page 26 shows the estimated survivor functions of the TIER 1 (high school graduate) population for cohort 84 broken down by mental group. The higher aptitude recruits display the greatest survivability in each cohort. Note that the behavior is similar to that illustrated in Figure 4 on page 22. The similar results are

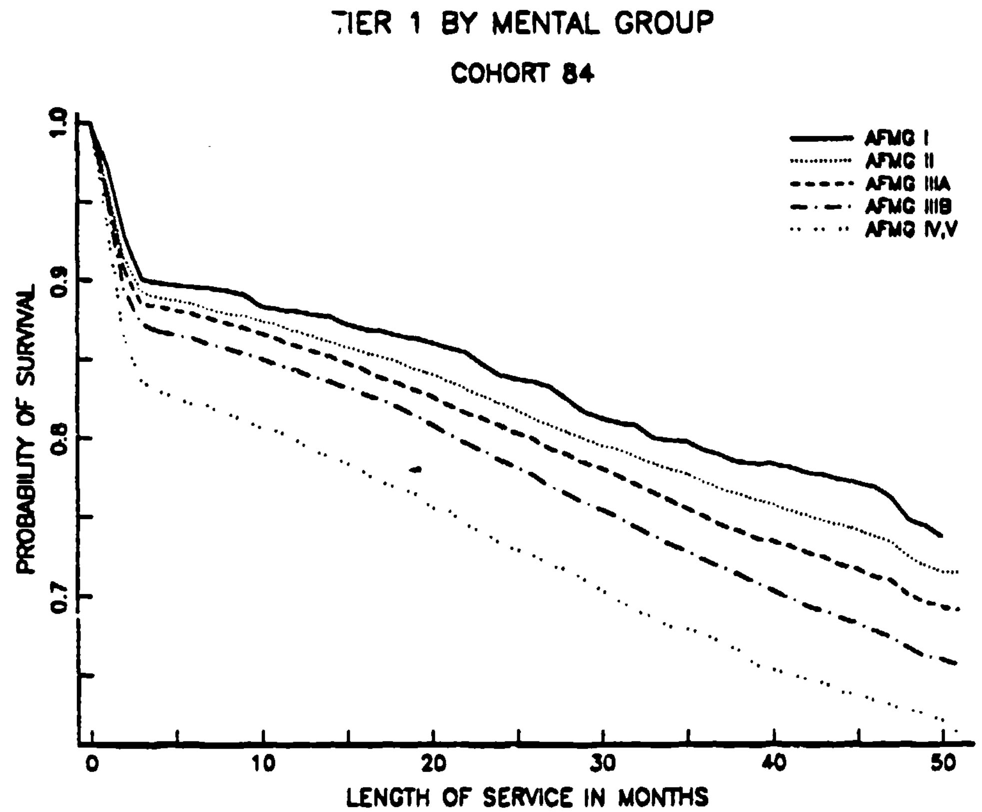

Figure 7. Survivor Functions, TIER 1 by AFMG, Cohort 84

hardly surprising as the TIER 1 population constitutes about ninety percent of the entire population. The remaining cohorts are displayed in Figure 29 to Figure 31 of Appendix B.

b. TIER 2

Figure 8 on page 28 shows the estimated survivor functions for the TIER 2 (alternate HS credential) population for cohort 84 broken down by mental group. Note that the TIER 2 survivor functions are not as well behaved as those for the TIER 1 population. This is likely a consequence of the smaller cell sizes. As Figure 8 demonstrates, however, the effect of mental group is, while perhaps a bit weaker, essentially the same for the TIER 2 recruits as for the TIER 1 population. The higher aptitude recruits again exhibit greater survivability in each cohort. Figure 32 on page 69 of Appendix B contains the TIER 2 by AFMG survivor functions for the remaining cohorts. While some mental groups are omitted due to an insufficient number of observations, those that are displayed demonstrate the same pattern of behavior as seen in Figure 8.

c. TIER 3

The TIER 3 (non-HS graduate) population, when broken down by mental group, does not follow the same pattern of behavior as the other two education groups. Figure 9 on page 29 and Figure 10 on page 30 show the estimated survivor functions for TIER 3 recruits by mental group for cohorts 84 through 87.

Note that the survivor functions for the different mental groups are relatively close together and there is a larger degree of randomness to their ordering. The higher aptitude TIER 3 recruits don't portray the consistently greater survivability characteristic of the other education groups. In fact, no discernable pattern of behavior is evident from Figure 9 and Figure 10. In this case, the graphical analyses fail to provide a clear picture of the relationship and empirical comparisons over the pooled sample must be relied upon. The empirical results of the next section show that the higher aptitude TIER 3 recruits do demonstrate greater survivability, when the data is combined across all cohorts. However, the relationship between aptitude and attrition is still rather weak among the TIER 3 population.

d. Empirical Comparison

Combining covariates reduces the effectiveness of the purely graphical analysis. In this case, there are fifteen subgroups formed when the population is partitioned by both DOD education code and armed forces mental group. In order to perform a meaningful evaluation of the various subgroups, the cohorts (84,85,86,87) were pooled into one large sample. Estimated survivor functions were then computed for each subgroup. The results of this analysis are contained in Table 5 on page 31.

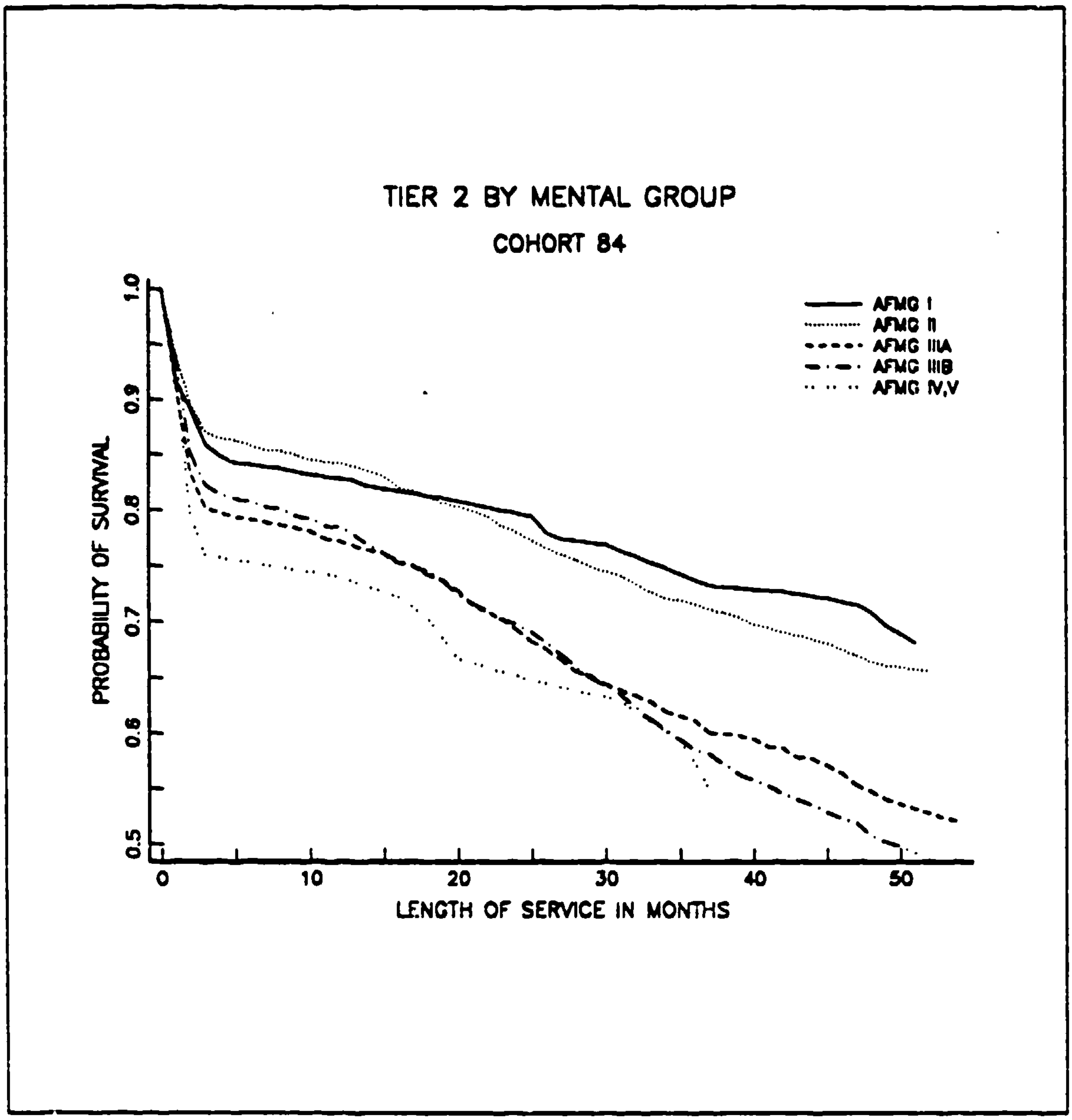

Figure 8. Survivor Functions, TIER 2 by AFMG, Cohort 84

The first column of Table 5 identifies the subgrouping, and the second column expresses the subgrouping size as a percentage of the total population. The next three columns represent the estimated survival probabilities for one, two, and three years. Finally, the last column assigns a rank to each subgroup based on its probability of survival to three years.

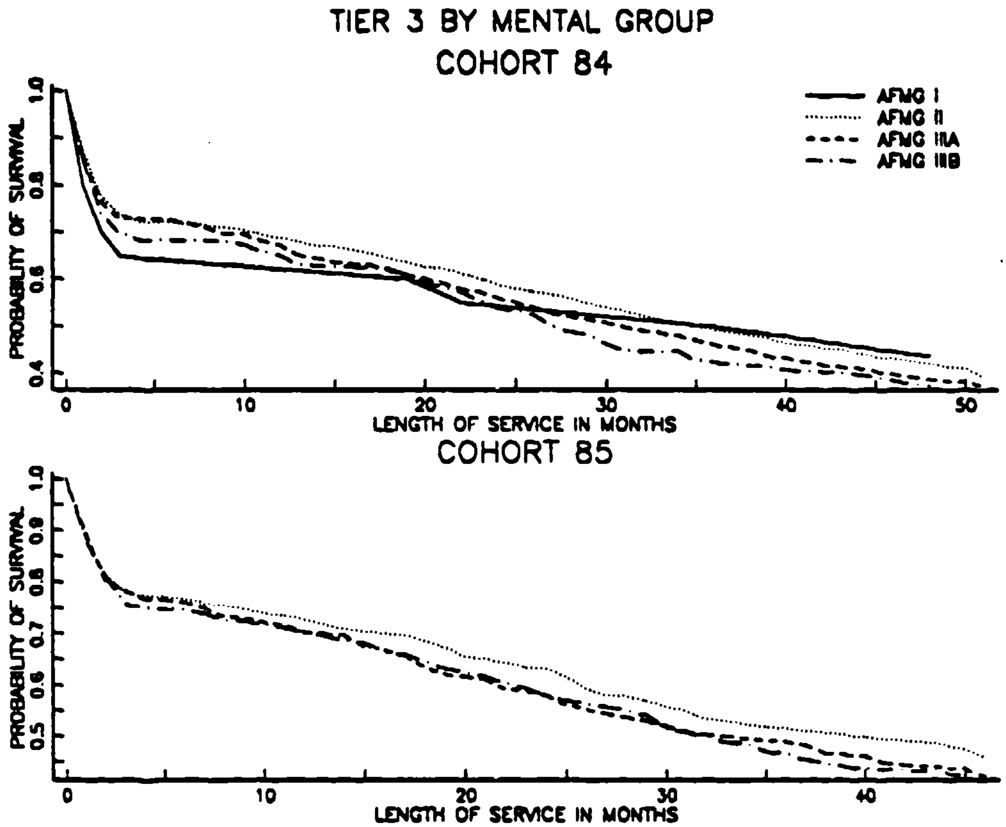

Figure 9. Survivor Functions, TIER 3 by AFMG, Cohorts 84 and 85

As expected, the subgroupings with the best rankings are the TIER 1 and high aptitude TIER 2 recruits. The most notable result is that the vast majority of *all* TIER 1 and TIER 2 recruits demonstrate better survivability than even the high aptitude TIER 3 recruits.

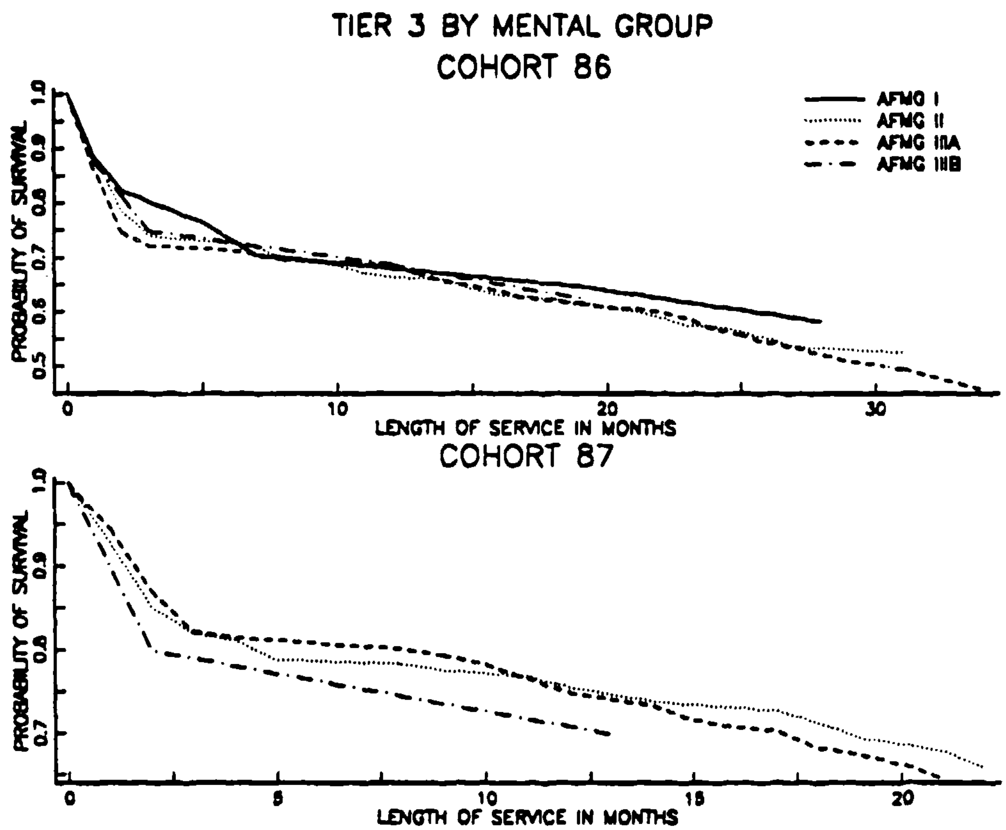

Figure 10. Survivor Functions, TIER 3 by AFMG, Cohorts 86 and 87

Table 5. SUMMARY OF EDUCATION AND AFMG GROUPINGS FOR COHORTS 84-87

Grouping	Frequency %	$\hat{S}(12 Months)$	$\hat{S}(24 Months)$	$\hat{S}(36 Months)$	Rank
TIER 1, AFMG I	2.41	.8791	.8288	.7784	1
TIER 1, AFMG II	28.09	.8669	.8165	.7260	2
TIER 1, AFMG IIIA	21.69	.8567	.7993	.7381	3
TIER 2, AFMG I	0.31	.8350	.7930	.7344	4
TIER 1, AFMG IIIB	33.81	.8421	.7831	.7158	5
TIER 2, AFMG II	2.65	.8383	.7719	.7063	6
TIER 1, AFMG IV,V	2.01	.8152	.7502	.6855	7
TIER 2, AFMG IIIA	2.07	.7914	.7092	.6183	8
TIER 3, AFMG I	0.04	.7155	.6594	.6102	9
TIER 2, AFMG IIIB	3.70	.7824	.7013	.6040	10
TIER 2, AFMG IV,V	0.10	.7705	.6642	.5879	11
TIER 3, AFMG II	1.09	.7036	.6047	.5054	12
TIER 3, AFMG IIIA	1.17	.6965	.5788	.4776	13
TIER 3, AFMG IIIB	0.46	.6831	.5647	.4516	14
TIER 3, AFMG IV,V	0.01	$<$ .5385	--	--	15

2. Education with Moral Waiver

The covariate moral waiver, when considered alone, proved to be a weak predictor of recruit suitability. The purpose of this section is to determine whether this relationship holds when the population is first partitioned by DOD education code.

a. TIER 1

Figure 11 on page 33 shows the estimated survivor functions of the TIER 1 population, pooled over cohorts 86 and 87, when broken down by the presence/non-presence of a moral waiver. As evidenced by the survivor functions, the recruits possessing a moral waiver display a higher attrition rate throughout their length of service. The curves in Figure 11 are similar to those in Figure 5 on page 24 and Figure 6 on page 25, hardly a surprising result considering that TIER 1 recruits constitute close to ninety-percent of the total population.

b. TIER 2

The estimated survivor functions for the TIER 2 population by presence/non-presence of a moral waiver are presented in Figure 12 on page 34. Note that, as before, the smaller cell sizes lead to more poorly behaved curves. However, it appears that the behavior is essentially the same as with the TIER 1 recruits in Figure 11 on page 33. In this case, though, the relationship between moral waiver and attrition doesn't seem to be as strong.

c. TIER 3

Figure 13 on page 35 contains the estimated survivor functions, by moral waiver, for the TIER 3 recruits of cohorts 86 and 87. It is unfortunate that more data from this particular subgrouping are not available to analyze, because Figure 13 shows some interesting behavior. Once again, the TIER 3 recruits behave differently than their TIER 1 and TIER 2 counterparts. The TIER 3 recruits with a moral waiver display a markedly greater attrition rate during the first three months of service. Of course, due to the small sample size, more observations are needed before formulating any solid conclusions.

d. Empirical Comparison

The recruit population for cohorts 86 through 88 was pooled and then partitioned into subgroups by DOD education code and presence/non-presence of a moral waiver. Each group was then evaluated in terms of survivability and ranked against the other groups. The results are contained in Table 6 on page 36. In this case, the ranking

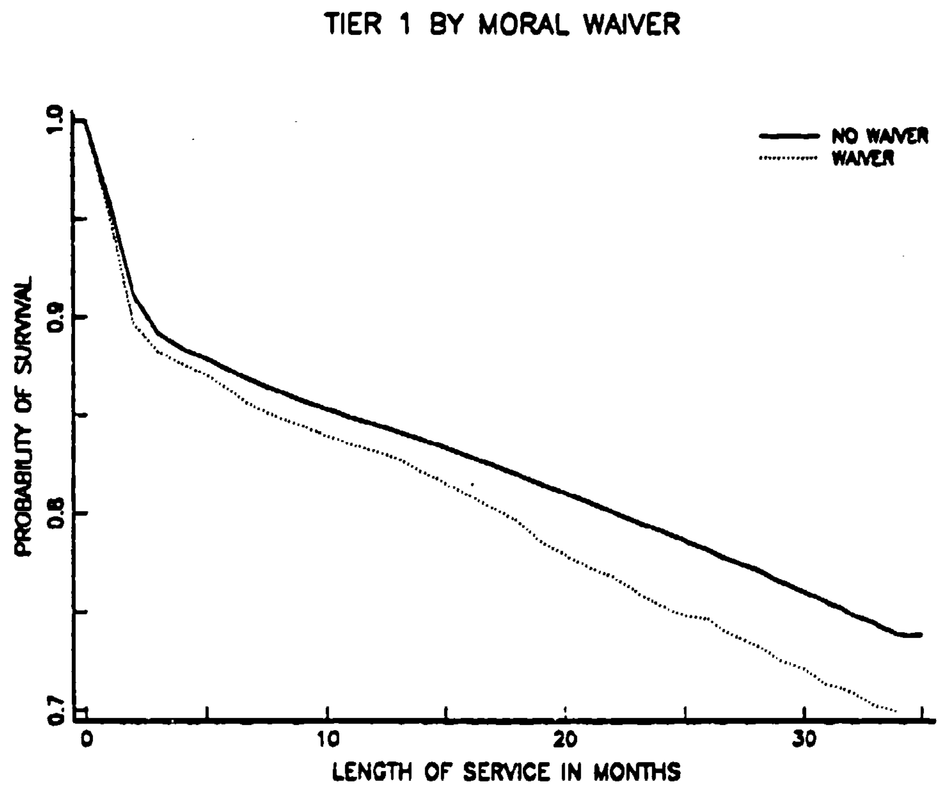

Figure 11. Survivor Functions, TIER 1 by Moral Waiver for Grouped Sample from Cohorts 86 and 87

criteria is probability of survival to 30 months. As Table 6 shows, the presence of a moral waiver fails to alter the ordering established by education level.

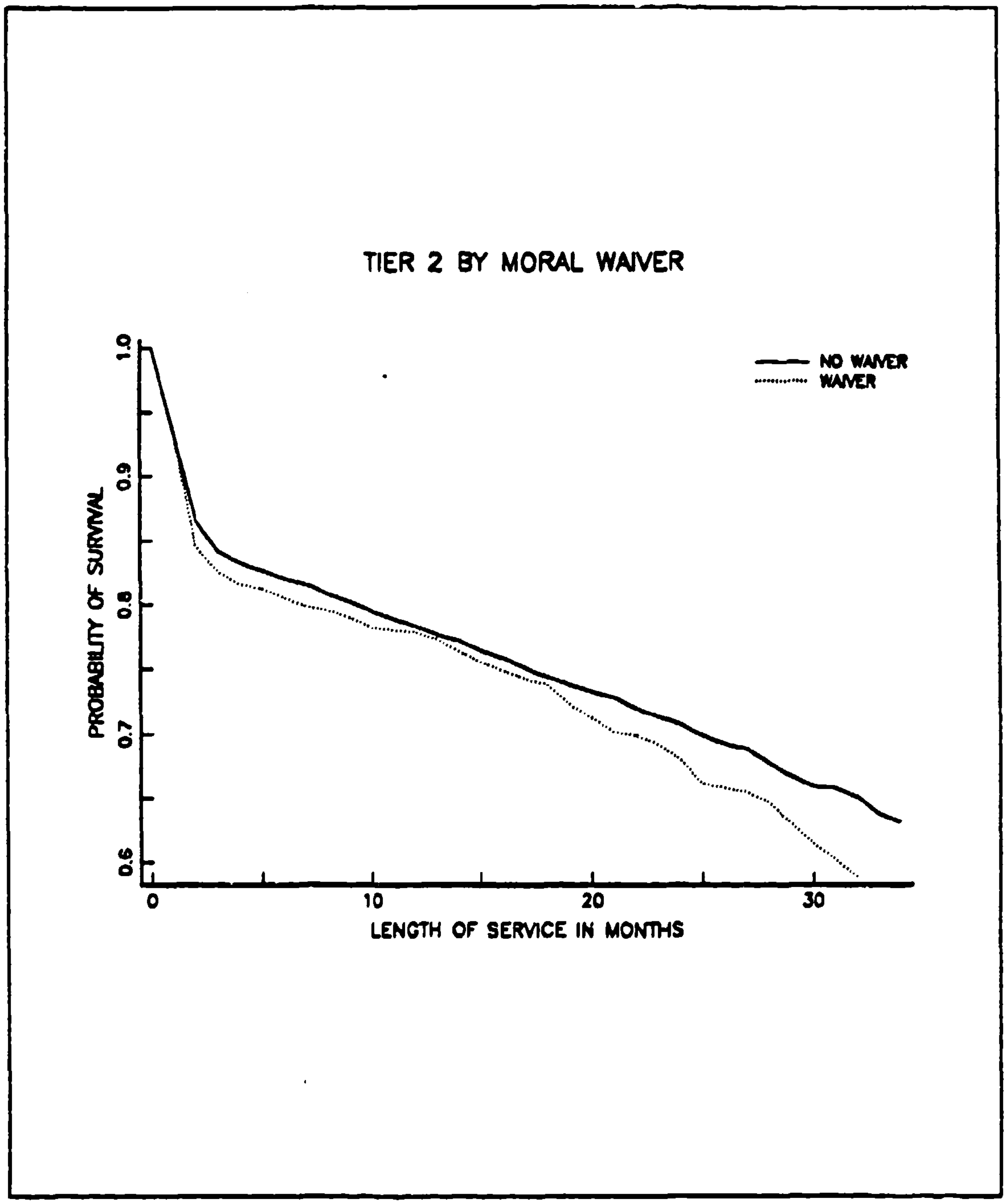

Figure 12. Survivor Functions, TIER 2 by Moral Waiver for Grouped Sample from Cohorts 86 and 87

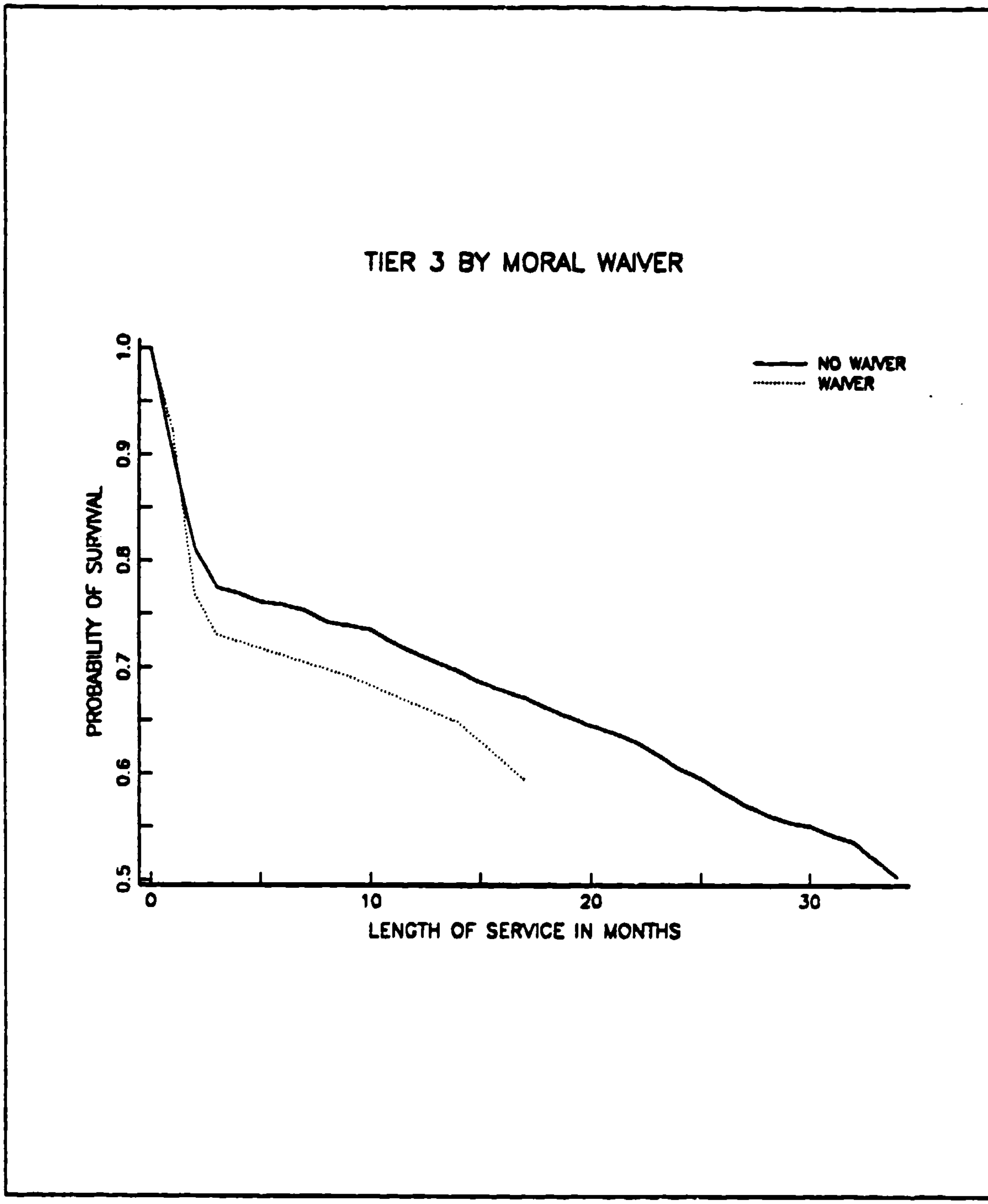

Figure 13. **Survivor Functions, TIER 3 by Moral Waiver for Grouped Sample from Cohorts 86 and 87**

Table 6. SUMMARY OF EDUCATION AND MORAL WAIVER GROUPINGS FOR COHORTS 86, 87

Grouping	Frequency %	$\hat{S}(12 Months)$	$\hat{S}(24 Months)$	$\hat{S}(30 Months)$	Rank
TIER 1 No Waiver	81.80	.8455	.7915	.7604	1
TIER 1 Waiver	7.50	.8320	.7531	.7208	2
TIER 2 No Waiver	7.16	.7837	.7073	.6599	3
TIER 2 Waiver	0.96	.7790	.6801	.6151	4
TIER 3 No Waiver	1.92	.7138	.6057	.5516	5
TIER 3 Waiver	0.04	< .6664	--	--	6

C. ALTERNATE HIGH SCHOOL CREDENTIALS

As shown in the *Single Covariate Effects* portion of this report, the attrition behavior of the TIER 2 population seems to be subject to a large degree of variability. In this section a closer look will be taken at the TIER 2 recruits, specifically the attrition behavior of each of the alternate high school credentials. The data from cohorts 84 through 87 are pooled in order to perform this analysis. Although this is not the preferred course of action, it becomes a necessary one in light of the relatively small number of TIER 2 recruits.

Table 7 gives the breakdown, by type of alternate high school credential for cohorts 84 through 87.

Table 7. FREQUENCY TABLE OF ALTERNATE HIGH SCHOOL CREDENTIALS FOR COHORTS 84 TO 87

Credential	Frequency
General Education Development (GED)	467 (4.16 %)
Correspondence School Certificate	90 (0.80 %)
High School Certificate of Attendance	7800 (69.52 %)
Occupational Program Certificate	2863 (25.52 %)

1. Credential Type

Figure 14 on page 38 shows the estimated survivor functions, by type of alternate high school credential, of the TIER 2 recruits for cohorts 84 through 87. From this plot, it is evident that large differences exist in the attrition behavior among different alternate credentials. At one extreme lie the occupational program certificate holders whom, in terms of attrition, essentially behave like TIER 1 recruits. At the other extreme are the GED holders, whose attrition behavior parallels that of TIER 3 recruits.

Aside from the obviously large differences in the attrition behavior among the alternate credentials, Figure 14 provides at least one other interesting insight. GED and correspondence school certificate holders display the greatest attrition of the four groups. These credentials also require the least amount of "seat-time" in school. Conversely, members of the other two groups, high school certificate of attendance and occupational program certificate holders are required to complete 11 to 12 years of formal school (See Appendix A for a more complete description of the requirements of each alternate credential). This seems to indicate a positive correlation between "seat-time"

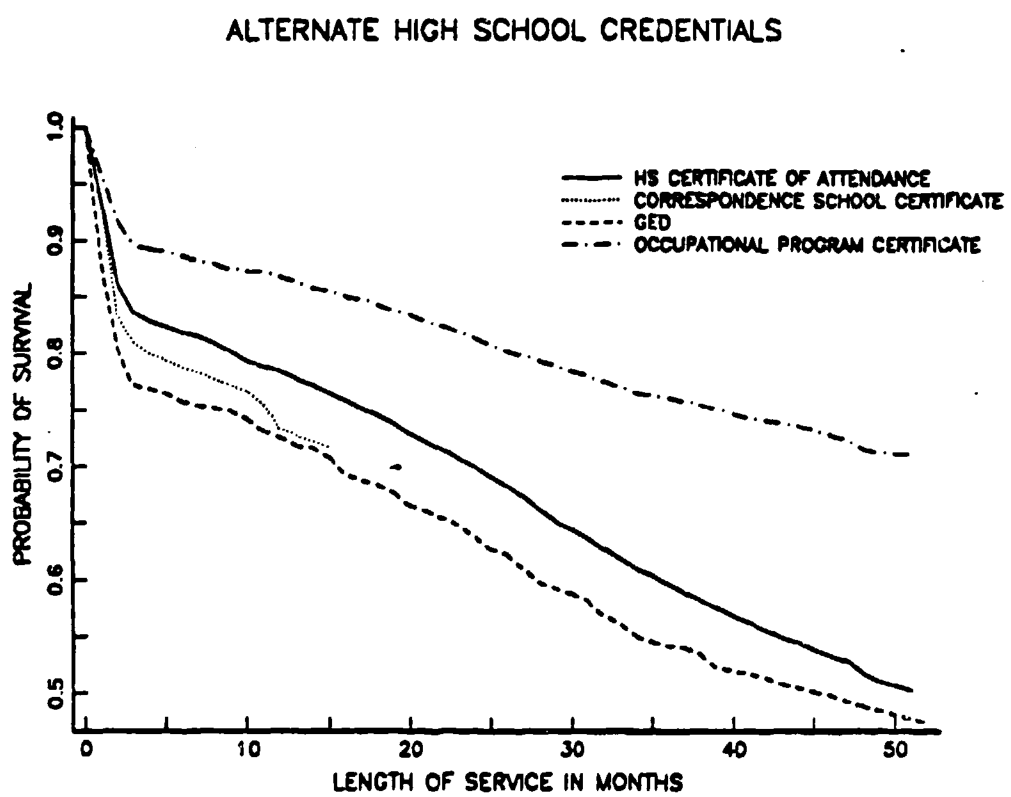

Figure 14. Survivor Functions by Type of Alternate High School Credential

and the probability of survival, which lends more credibility to the hypothesis that "seat-time" in school, and thus amount of social interaction, is the reason that high school diploma graduates are, as a group, much better suited for military service.

2. Credential Type with Aptitude

In the *Combined Covariate Effects* portion of this report it was established that the TIER 2 population, when partitioned by mental group, demonstrates a positive relationship between aptitude and survivability. Figure 8 on page 28 and Figure 32 on page 69 of Appendix B illustrate this relationship.

In examining the different alternate credentials, it is of interest to determine if aptitude level has the same effect on each in terms of attrition behavior. Figure 15 on page 41 shows the estimated survivor functions of the pooled GED holders, by mental group. Note that the established inverse correlation between aptitude level and attrition is evident. This is an interesting result because to this point the attrition behavior of the GED holders has appeared to parallel that of the TIER 3 population. However, recall from Figure 9 on page 29 and Figure 10 on page 30 that the relationship between aptitude and attrition is considerably weaker for the TIER 3 population. Therefore, despite the similarities of their survivor functions, the GED holders and TIER 3 recruits have at least one major difference: A high aptitude GED holder is much more likely to complete his first term of service than a medium or low aptitude GED holder, but a high aptitude TIER 3 recruit is probably not much more likely to complete his first term of service than a medium or low aptitude TIER 3 recruit.

Figure 16 on page 42 shows the estimated survivor functions by mental group of the pooled occupational program certificate and high school certificate of attendance holders. There is nothing notable in these plots, with the exception of the odd attrition behavior on the part of the members of AFMG's I,IV, and V. However, in the case of these mental groups, small sample sizes are the most logical explanation for the erratic attrition behavior. Otherwise, the relationship between aptitude and attrition is evident, especially among the occupational program certificate holders. This is further evidence that the graduates of occupational programs demonstrate a pattern of attrition behavior quite similar to that of high school diploma graduates (Recall from previous discussion that the relationship between aptitude and attrition was strongest among the TIER 1 recruits).

D. POST BOOT CAMP ATTRITION

This section deals with the analyses of covariate effects on survivability beyond the four month point of service. A potentially useful practice in performing analyses of Marine enlisted attrition is to examine separately the first four months of service and the time in service beyond four months. Justification for proceeding in this manner emerges from the estimated survivor functions computed to this point. It is evident that the rate of attrition during the first few months of service is much higher, and that the primary source of this attrition is recruit failure from boot camp. As a matter of convenience, the following labels will be used to describe the two categories of recruit attrition:

- **Boot Camp Attrition**-attrition that occurs during the first four months of service.
- **Post Boot Camp Attrition**-attrition that occurs after the first four months of service.

A question immediately arises as to whether the effects of the covariates on attrition behavior are consistent for the two periods. As an example, consider that attrition-wise, some recruits are high risk during boot camp, but relatively low risk after boot camp. This is not an unreasonable hypothesis considering that boot camp is a unique environment and that many of the experiences of boot camp are unlikely to be encountered elsewhere in the Marine Corps.

The examination of boot camp attrition is straightforward, merely requiring the inspection of the first four months of the survivor functions presented earlier in this chapter for the entire length of service. Since the survivor functions have a common starting point, comparisons of covariate effects during boot camp are direct. What is not always clear, however, are the covariate effects after the survivor functions level off at around the three to four month point. It appears by visual inspection that the differences in attrition among the covariate levels are not as pronounced after the four month point as they were during boot camp. However, since there is no common reference point on the survivor functions at the four month point, meaningful comparisions are difficult to make.

Conditioning the estimated survivor functions on the completion of four months of service provides a common reference point that aids in the analyses of post boot camp attrition. The *conditioned product-limit estimate* of the survivor function, $\hat{S}_c(t)$, is computed as follows:

$$\hat{S}_c(t) = \frac{\hat{S}(t)}{\hat{S}(4)} \quad ; \; t \geq 4.$$

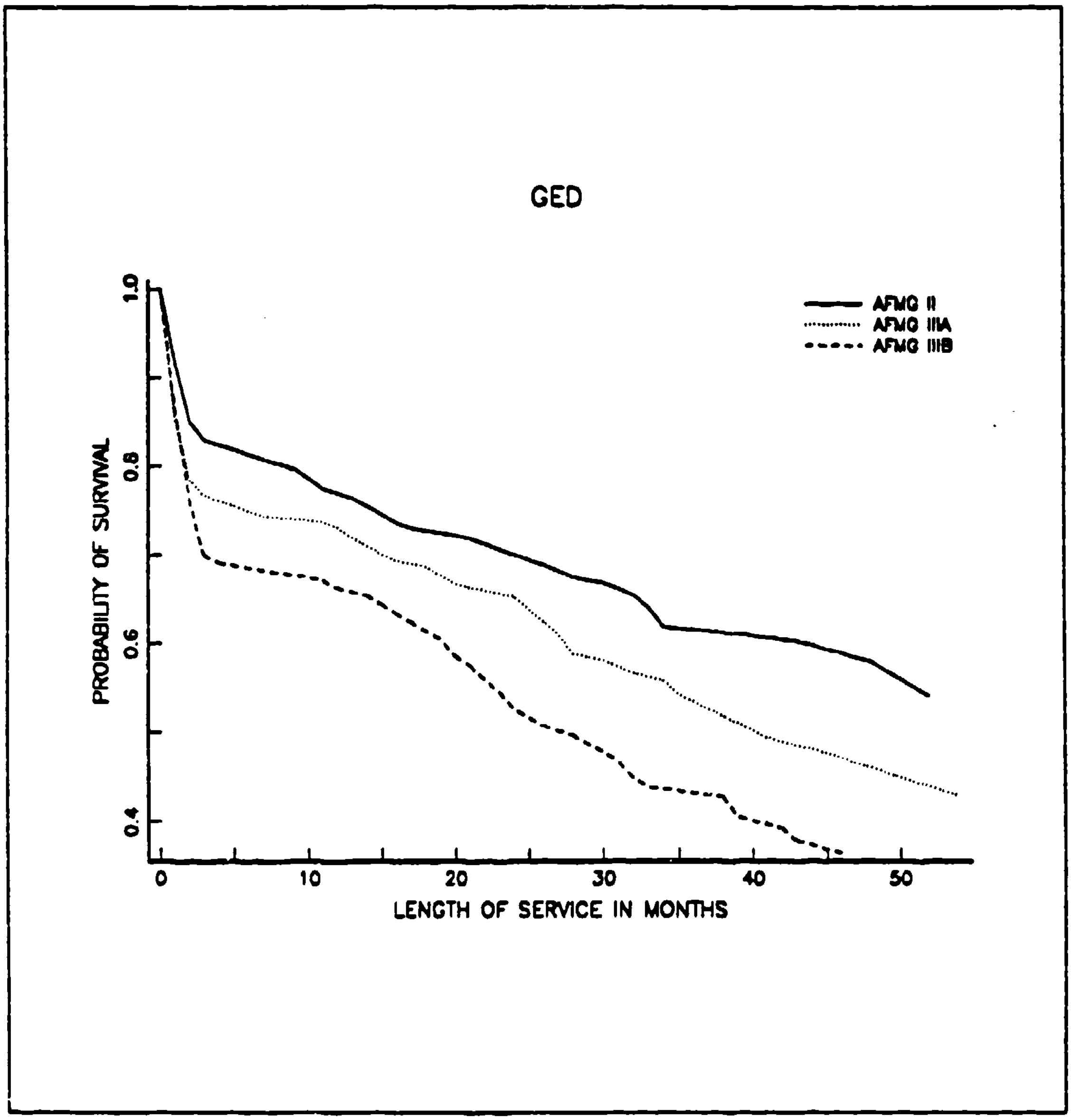

Figure 15. Survivor Functions by Mental Group of GED Holders

1. Single Covariate Effects

a. Education

Figure 17 on page 43 shows the conditioned survivor functions, by DOD education code, for cohort 84. It is evident from the plot that education level is an effective predictor of attrition behavior throughout the entire length of service. Since the other cohorts display similar behavior, they are not included.

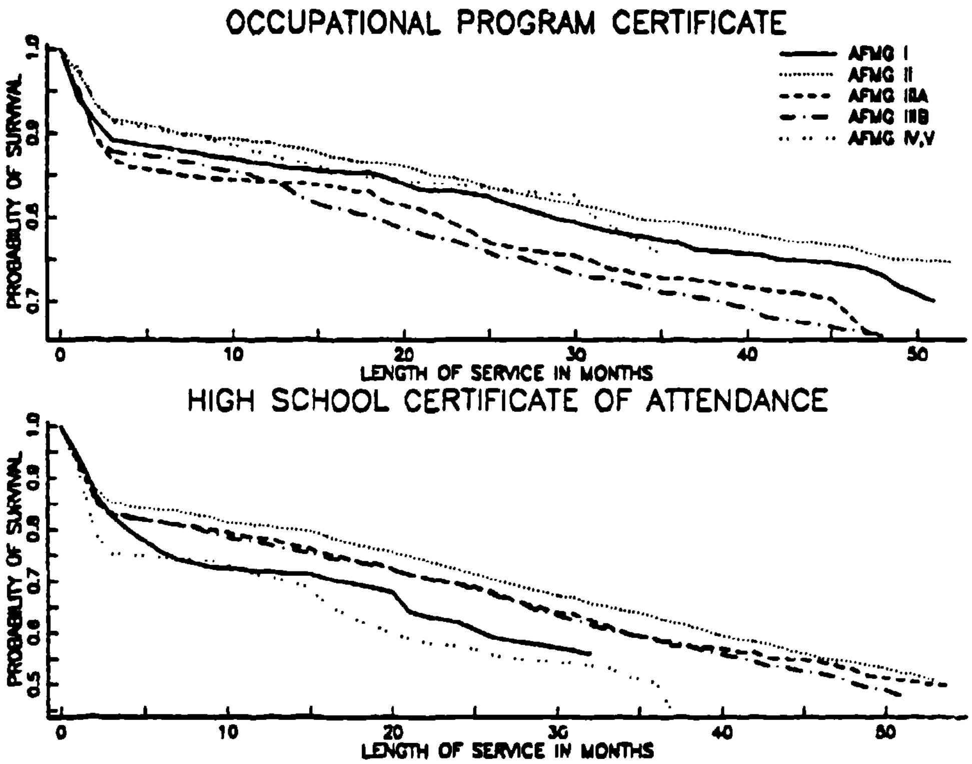

Figure 16. Survivor Functions by Mental Group of Occupational Program Certificate and High School Certificate of Attendance Holders

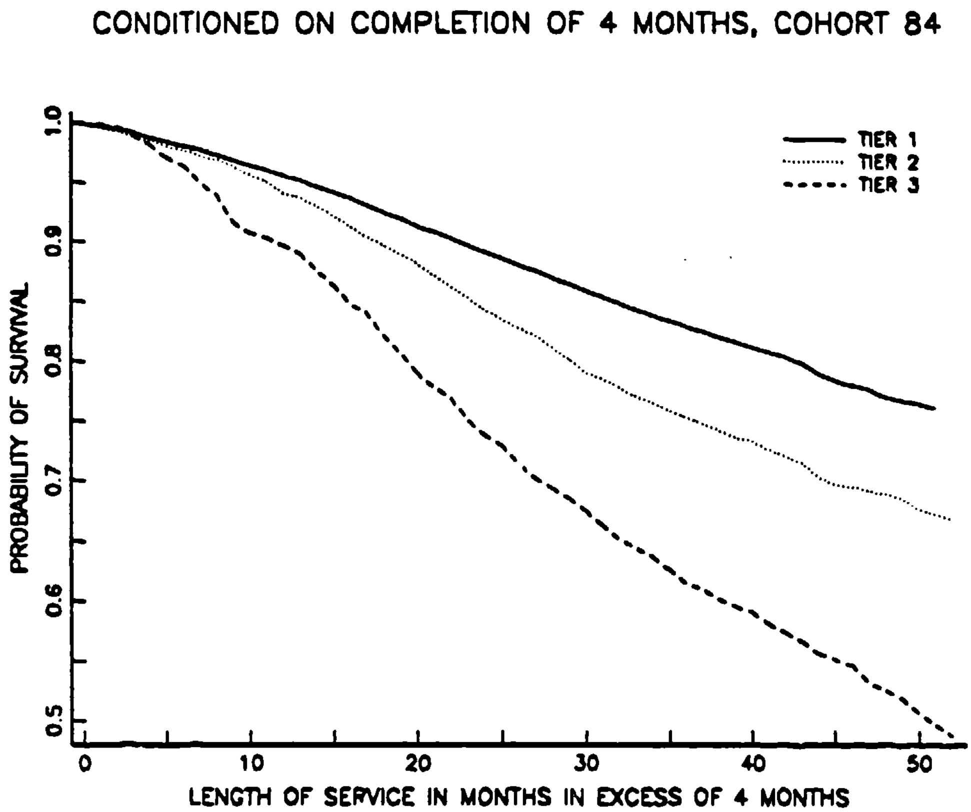

Figure 17. Conditioned Survivor Functions by DOD Education Code, Cohort 84

b. Mental Group

Figure 18 on page 45 and Figure 19 on page 46 show the conditioned survivor functions, by Armed Forces Mental Group, for cohorts 84 through 88. These plots reveal a trend present in the data. Consider Figure 18. Note that the relationship between aptitude and attrition holds for cohort 84. The relationship holds as well for cohort 85, but is weaker as evidenced by the tighter grouping of the survivor functions.

Now consider Figure 19. The relationship between aptitude and attrition continues to weaken as the cohorts become more recent. In fact it would be difficult to make any concrete statement concerning the effects of aptitude on post boot camp attrition based on the three plots of Figure 19.

This finding helps to explain the tightening of the mental group survivor functions noted earlier in this chapter. Aptitude remains an effective predictor of boot camp attrition. A visual inspection of Figure 4 on page 22 and Figure 25 to Figure 28 of Appendix B provides confirmation. However, it appears that the higher aptitude enlistees that survive boot camp are no longer much more likely to complete their service obligation than lower aptitude boot camp graduates. Whether this is an indicator of social change or simply a local phenomena within the data set is a question for further study. It is not implausible that some change in policy enacted during the period contributed to the change in attrition behavior.

c. *Moral Waiver*

Figure 20 on page 47 shows the conditioned survivor functions, by moral waiver, for cohorts 86, 87, and 88. The effect of moral waivers seems to remain consistent for boot camp and post boot camp.

The issue of moral waivers are often raised in conjunction with the screening of personnel for sensitive positions. The assignment to such positions normally occurs at the end of boot camp. Other issues aside concerning the reliability of enlistees with moral waivers, from the standpoint of probability of attrition, the plots of Figure 20 support the use of moral waivers in the screening process.

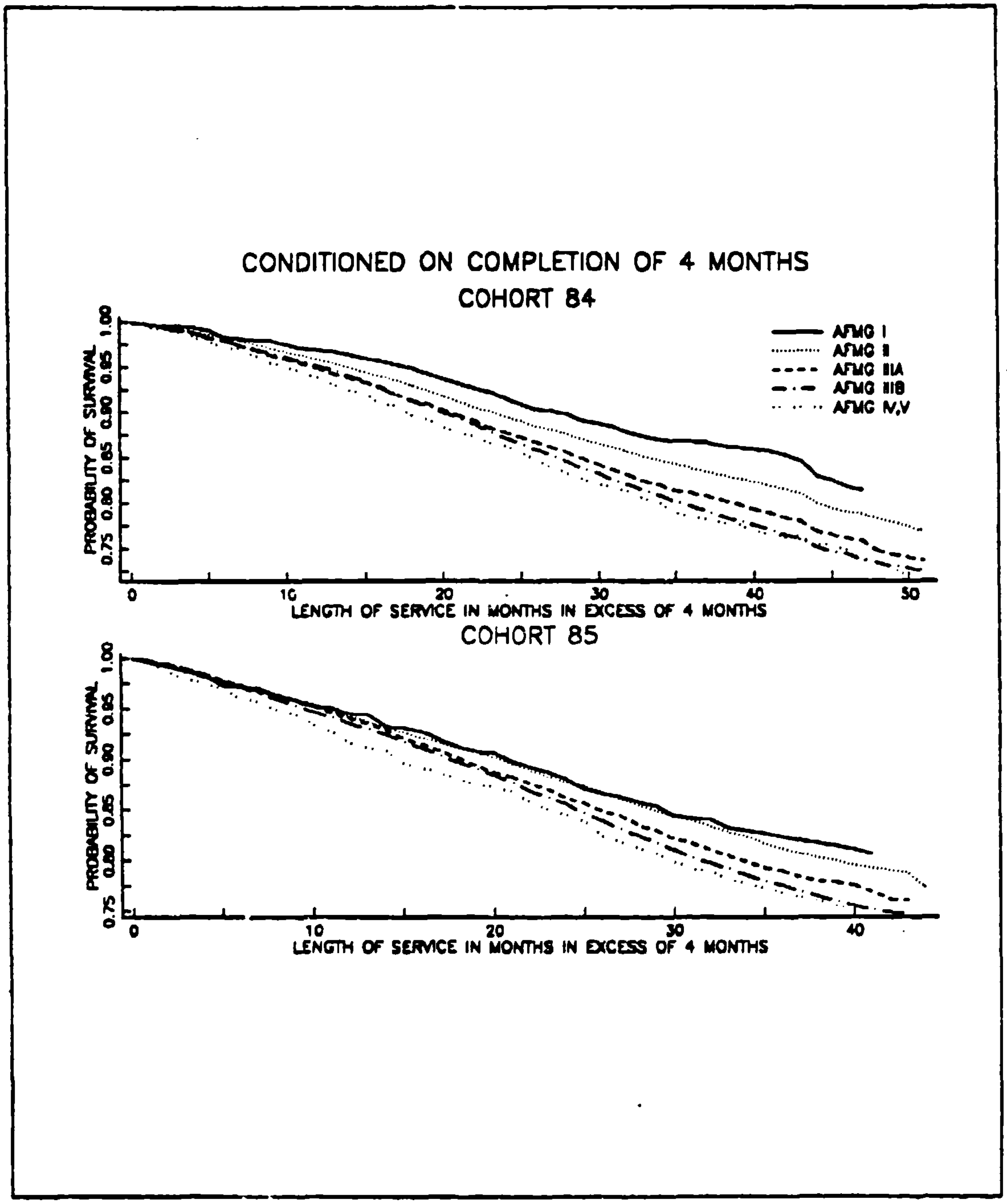

Figure 18. Conditioned Survivor Functions by Mental Group, Cohorts 84 and 85

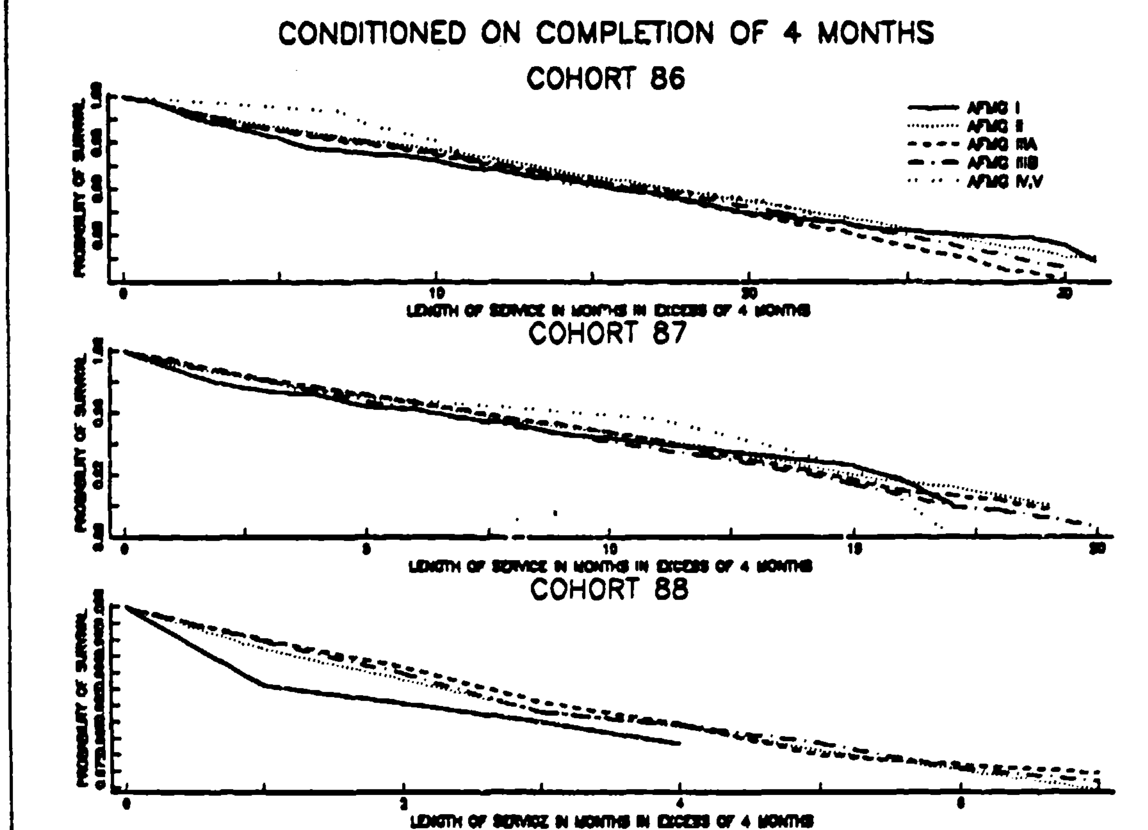

Figure 19. Conditioned Survivor Functions by Mental Group, Cohorts 86 to 88

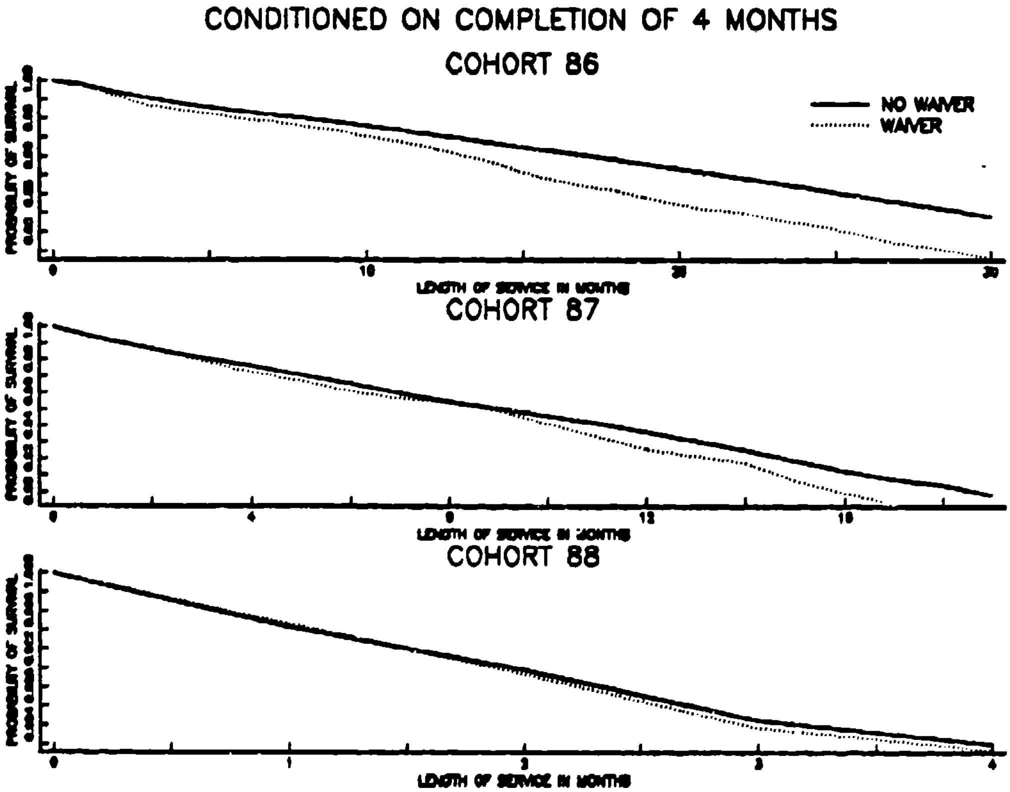

Figure 20. Conditioned Survivor Functions by Moral Waiver, Cohorts 86 to 88

VI. ATTRITION MODEL

The *product-limit estimate* of the survivor function is an effective tool for analyzing attrition behavior. However, the computational effort required is tremendous, and the covariate values must be categorical (or converted to categorical as in the case of AFQT score). An accurate attrition model could probably lead one to the same conclusions that were reached in the *Survival Analysis* portion of this report. However, a model would typically be many times more computationally efficient, and could easily accommodate non-categorical data as well as the addition of new covariates. With these considerations in mind, this chapter embarks on the development of such a model.

The chapter is broken down into two sections. The first deals with the technical details of the model, the second demonstrates the model in use and provides an evaluation of effectiveness.

A. METHODOLOGY

This section provides an overview of the development of the model. The primary source used in developing the model is Kalbfleisch and Prentice [Ref. 6: pp. 50-55]. Much of the notation and methods of presentation used in this section closely parallels that used in their text. Readers in pursuit of a more detailed development of the log-linear model are referred there.

1. Weibull Distribution

The Weibull distribution is an important generalization of the exponential distribution. The probability density function of the Weibull is given by

$$f(t) = \lambda p(\lambda t)^{p-1} e^{-(\lambda t)^p},$$

and the survivor function is

$$S(t) = e^{-(\lambda t)^p}.$$

Thus the Weibull is a two parameter distribution that, unlike the exponential, allows for a power dependence of the hazard on time. The hazard function for the Weibull is given by

$$\lambda(t) = \lambda p(\lambda t)^{p-1}.$$

The additional parameter p makes the Weibull much less restrictive than the exponential. In fact, the exponential distribution is simply a Weibull with the shape parameter (p) restricted to 1.

2. Log-Linear Regression Model

The Weibull distribution can be generalized to obtain a regression model by allowing the failure rate to be a function of the vector of covariates X. If the conditional (on X) hazard is

$$\lambda(t \mid X) = \lambda p(\lambda t)^{p-1} e^{X\beta},$$

then the conditional density of T is

$$f(t \mid X) = \lambda p(\lambda t)^{p-1} e^{X\beta} e^{-[(\lambda t)^{p} e^{X\beta}]}, \tag{3}$$

where $\beta = (\beta_1, \ldots, \beta_s)$ is a vector of regression parameters. In terms of $Y = \log T$, the model (3) is the linear model

$$Y = \alpha + X\beta' + \sigma W \tag{4}$$

where $\alpha = -\log \lambda$, $\sigma = p^{-1}$, $\beta' = -\sigma\beta$ and W has the extreme value distribution.

The p.d.f. for $Y = \log T$ can be written $\sigma^{-1} f(w)$, where $w = \dfrac{(y - X\beta)}{\sigma}$ and $X = (X_1, \ldots, X_s)$ is a regression vector corresponding to failure time t. In the case of the Weibull distribution $f(w) = e^{(w - e^{w})}$.

Suppose that there are n observations from the model (4). Allow δ_i to be an indicator variable that takes on the value of 1 if the observation is not censored and the value 0 if the observation is censored. If $S(.)$ is the survivor distribution for the Weibull and y_i is the logarithm of the observed survival time for the ith individual, the likelihood function may be written

$$L(\beta, \sigma, p) = \prod_{i=1}^{n} [\sigma^{-1} f(w_i)]^{\delta_i} S(w_i)^{1-\delta_i},$$

where $w_i = \dfrac{(y_i - X_i\beta)}{\sigma}$.

The terms of the gradient $\dfrac{\partial L}{\partial p}$, $\dfrac{\partial L}{\partial \sigma}$, and $\dfrac{\partial L}{\partial \beta_j}$, $(j = 1, \ldots, s)$ are then solved numerically to obtain estimates for p, σ, and β_j.

a. Accelerated Failure Time Model

The Weibull regression model is a log-linear model, which implies that the covariates act multiplicatively on the time scale (or additively on Y in (4) above). This class of model is known as the accelerated failure time model and is characterized by the following scaling of the failure time

$$T = e^{-X\beta} T_0.$$

Where T_0 is a failure time from a baseline distribution with values corresponding to zero for the covariates.

B. APPLICATION OF MODEL

1. Two-Piece Weibull

As previously noted, the survivor function for the Weibull distribution is

$$S(t) = e^{-(\lambda t)^p}.$$

It follows that

$$\log[-\log S(t)] = p(\log t + \log \lambda).$$

As a result, a convenient check for the Weibull distribution is provided by a plot of $\log[-\log \hat{S}(t)]$ versus $\log t$, where $\hat{S}(t)$ is the *product-limit estimate* of the survivor function. If the Weibull distribution is appropriate, the plot should give approximately a straight line.

Figure 21 on page 51 shows the plots of $\log[-\log \hat{S}(t)]$ versus $\log t$ for boot camp, and for post boot camp, cohort 86. As before, "boot camp" refers to the first four months of service and "post boot camp" to the conditioned survivor function given survival to at least four months. Since the plots are reasonably straight lines in both cases, the distribution of the entire survivor function may be thought of as a two-piece Weibull. That is the survivor function for the first four months of service is a Weibull with parameters λ_1 and p_1. The conditioned survivor function for length of service beyond four months becomes a Weibull with parameters λ_2 and p_2.

While the two-piece Weibull is not an easy distribution to arithmetically manipulate, it does provide a means by which to model recruit attrition. The log-linear regression model with the Weibull as the underlying distribution that was developed previously can be applied to the two separate "pieces" of the survivor function.

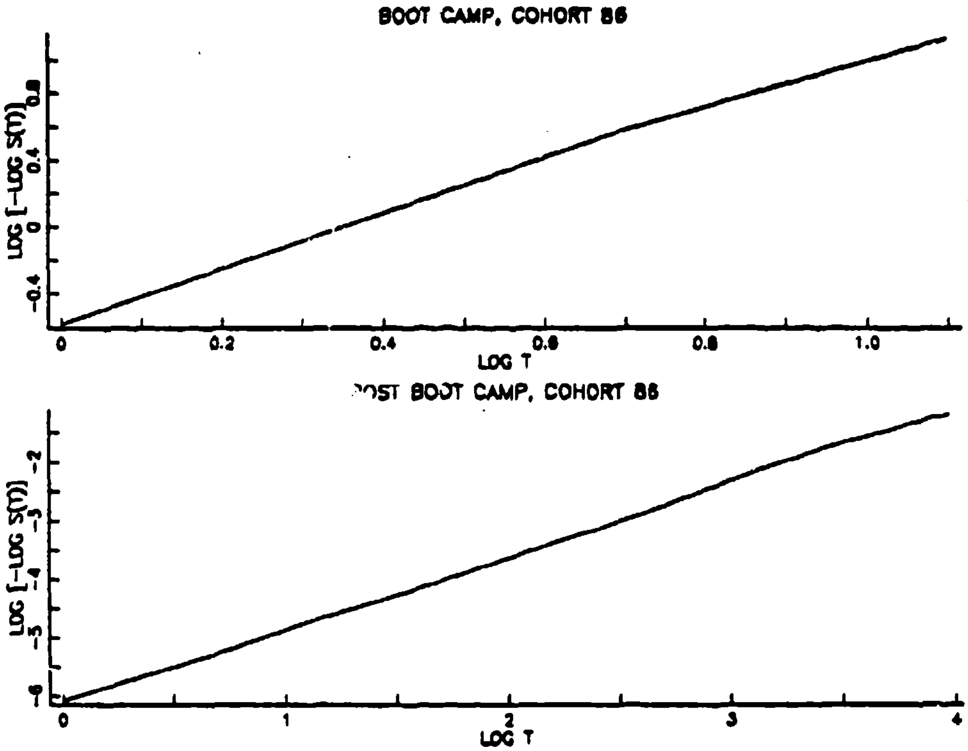

Figure 21. Fit of Cohort 86 to Weibull Distribution

The adequacy of the fit provided by the log-linear model depends greatly on the behavior of the survivor functions at the various covariate levels. If the survivor functions do not fit the Weibull, then the model will provide poor results. It can be shown using plots like those in Figure 21 on page 51 that the survivor functions at the different covariate values do fit the Weibull reasonably well.

2. Illustration of Model in Use

This section demonstrates the use of the log-linear regression model. Two things are gained through this exercise. First, a demonstration of the model in practice provides the best means by which to assess its capabilities and limitations. Second, the adequacy of the model can be evaluated by plotting the model survivor function versus the *product-limit estimate* of the survivor function.

For ease of presentation, the results from the two regressions performed are tabulated in the body of the text. The actual regression output generated by the SAS LIFEREG procedure is contained in Appendix D [Ref. 10: pp. 507-528].

The data from cohort 86 are used to perform the log-linear regression of length of service on the independent covariates of DOD education code, AFQT score, and moral waiver. The actual AFQT percentile is used instead of Armed Forces Mental Group. Thus, there are five independent variables (covariates) in the model. The variables, X_{TIER1}, X_{TIER2}, X_{TIER3} take on the value of 1 if the observation falls in that education category, 0 otherwise. Similarily, the variable X_{WAIVER} is 1 if a moral waiver was granted and 0 otherwise. The variable X_{SCORE} is a numeric value from 1 to 99.

Table 8 on page 53 contains the results of the regressions. The value in parenthesis next to the regression coefficient β_j is the significance level for rejecting the null hypothesis, H_0: $\beta_j = 0$. The middle column contains the parameter estimates and regression coefficients for boot camp, and the right column the same for post boot camp. The survivor functions for the two pieces of the service life may be solved for directly using

$$\tilde{S}(t) = e^{-(\lambda e^{-X\beta} t)^{\gamma}}, \tag{5}$$

where $\tilde{S}(t)$ is the modelled survivor function. The results from Table 8 together with the the covariate values X, provide for the solution of (5). Coupled with a knowledge of the probability of survival beyond four months, the covariate values are sufficient to model the survivor function for the entire length of service.

Table 8. REGRESSION RESULTS, COHORT 86

Parameters	Boot Camp	Post Boot Camp
λ (scale)	0.0204	0.0921
p (shape)	1.8438	1.1317
Regression Coefficients		
β_{TIER1}	-0.0247 (.6067)	2.3868 (.0001)
β_{TIER2}	-0.0568 (.2962)	2.0595 (.0001)
β_{TIER3}	-0.1710 (.0058)	1.7045 (.0001)
β_{SCORE}	-0.0006 (.2264)	0.0023 (.0041)
β_{WAIVER}	-0.1081 (.0004)	-0.2715 (.0001)

a. Example 1

In this case we wish to model the survivor function of recruits that are high school diploma graduates, belong in mental group IIIA, and possess a moral waiver. Thus the covariate values are

$$\begin{aligned} X_{TIER1} &= 1 \\ X_{TIER2} &= 0 \\ X_{TIER3} &= 0 \\ X_{SCORE} &= 57 \\ X_{WAIVER} &= 1. \end{aligned}$$

X_{SCORE} is assigned the value of 57, because it is the middle value of the range of scores for mental group IIIA.

Substituting the appropriate values from Table 8 along with the covariate values above into (5), a modelled survivor function for the entire length of service may be obtained. Figure 22 on page 54 shows this modelled survivor function plotted with the *product-limit estimate* of TIER 1 recruits from AFMG IIIA that possess a moral waiver. Note that after the five month point the fit is less than exact. However, since the actual survivor function does not appear to be well behaved, a precise fit is probably unattainable regardless of the modelling technique employed.

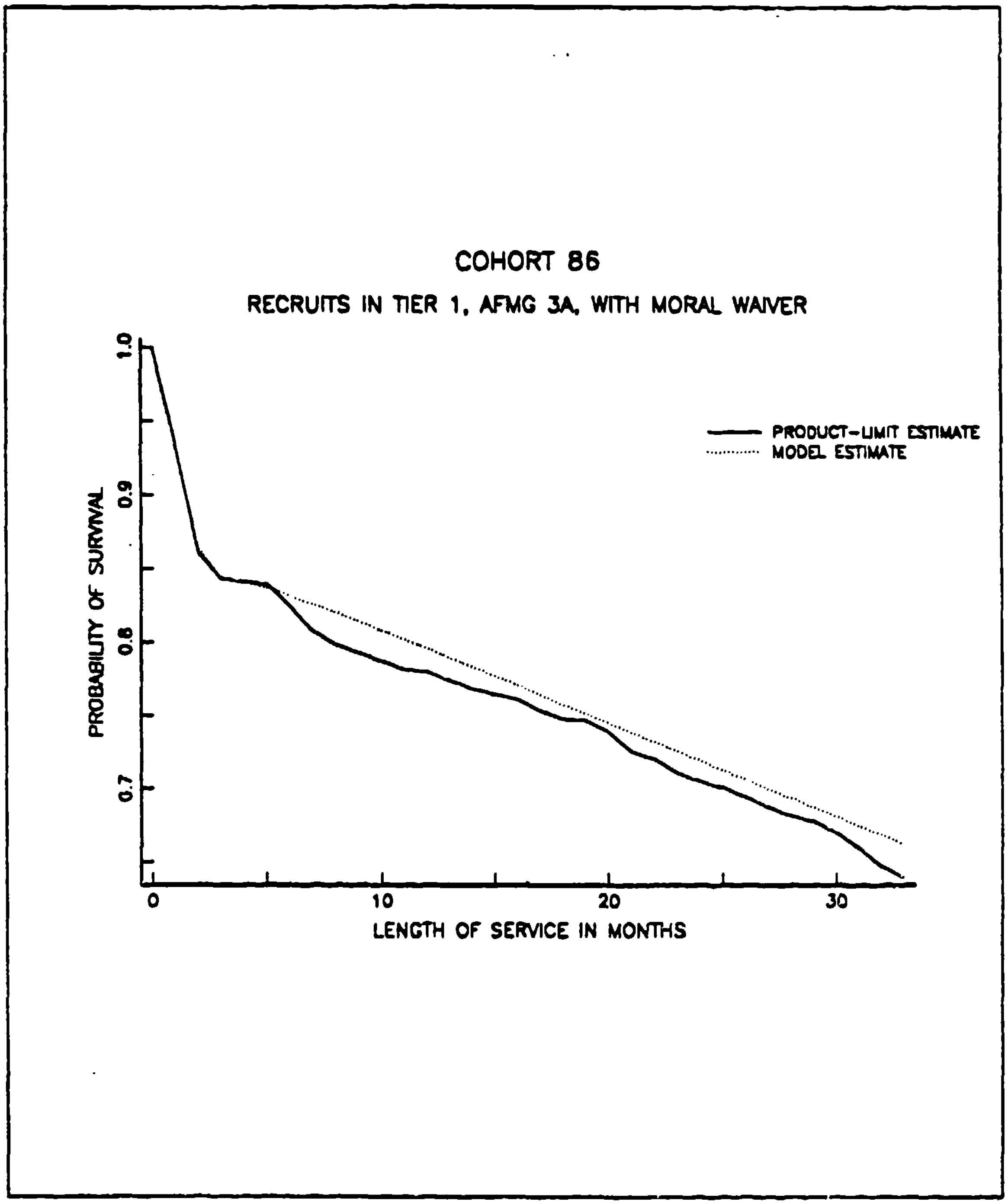

Figure 22. Comparision of Survivor Functions for Example 1

b. Example 2

In this case we wish to model the survivor function of recruits that have an alternate high school credential, belong in mental group IIIA, and do not possess a moral waiver. Thus the covariate values are

$$
\begin{aligned}
X_{TIER1} &= 0 \\
X_{TIER2} &= 1 \\
X_{TIER3} &= 0 \\
X_{SCORE} &= 57 \\
X_{WAIVER} &= 0.
\end{aligned}
$$

Figure 23 on page 56 shows the modelled survivor function plotted with the *product-limit estimate* of recruits with identical covariate values. The fit in this case is quite good.

3. Evaluation of Model

The log-linear regression model provides an adequate representation of the actual process. The two examples performed point out a major weakness of the model in that the quality of the fit obtained is dependent upon the behavior of the survivor function of the population (as determined by the covariate values) involved. The general rule for this dataset is that the larger the population subgrouping, the closer its survivor function approximates the Weibull. In this case, the most accurate fits were provided by the model when only one covariate was considered. Thus, while the model easily facilitates the introduction of additional covariates at minimal computational expense, the price is a poorer fit as a consequence of the smaller population sizes.

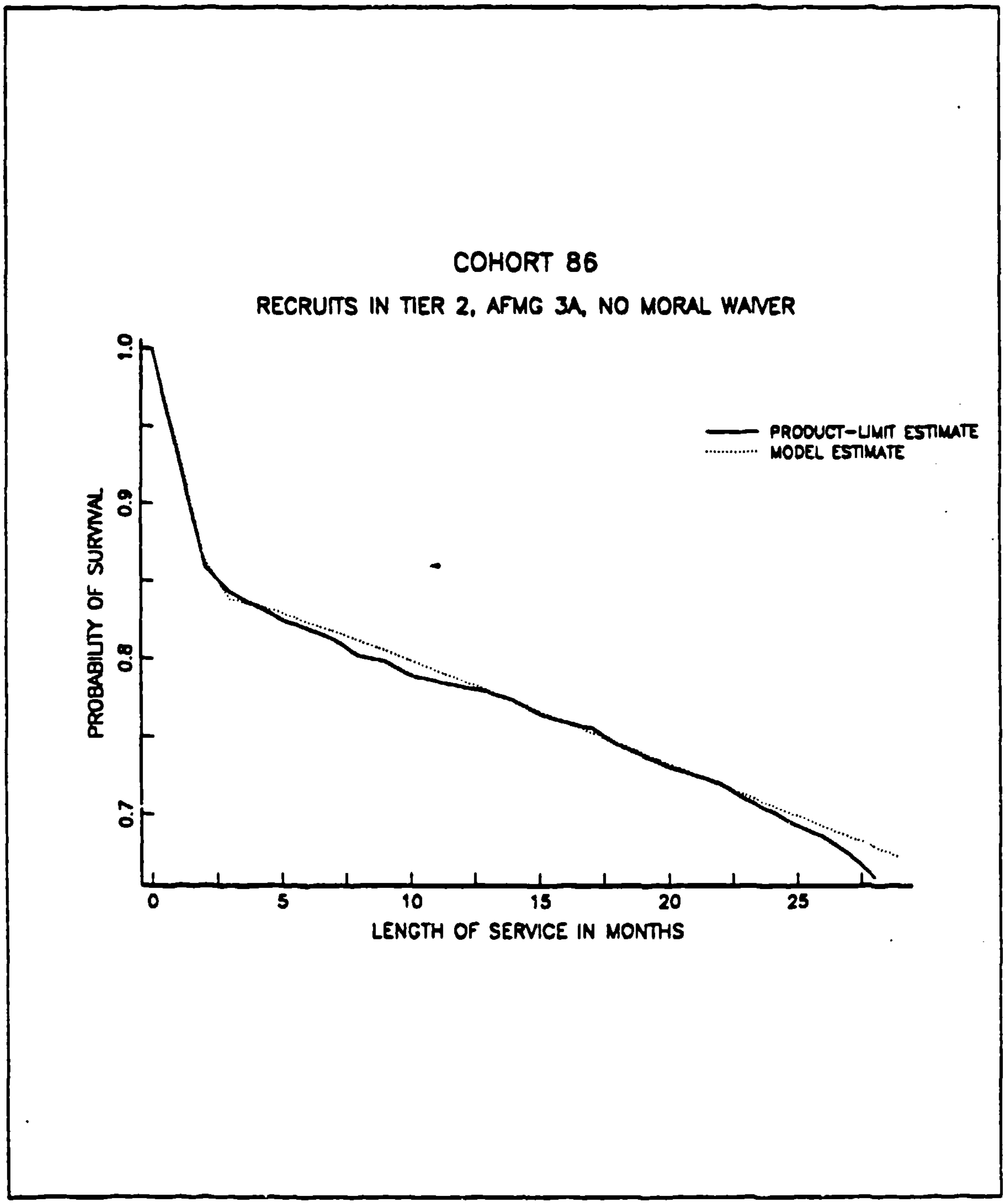

Figure 23. Comparision of Survivor Functions for Example 2

VII. SUMMARY AND CONCLUSIONS

A. SURVIVAL ANALYSIS

The *product-limit estimate* of the survivor function is an excellent tool to analyze the attrition behavior of a population when censored observations are present. The censoring mechanism accommodates special cases of attrition that require a different classification than non-EAS attrition without biasing the estimates of the survival probabilities. Examples of these special cases include the three-year service obligation and early release programs. The graphical plot of the *product-limit estimate* provides a more pleasing illustration than the usual table of survival probabilites associated with attrition studies. Additionally, interpretation of the plots doesn't require a background in statistics.

The survival analysis performed supported many of the traditional findings of military attrition studies and provided some interesting insights concerning alternate high school credential holders and on the relationship between aptitude and attrition. The most significant results are briefly outlined below.

- **Education**- Of the three covariates, education level based on DOD education code has the strongest relationship with attrition. The TIER 2 population is heterogeneous, as the attrition behavior varied significantly among the four types of alternate high school credentials. There is also indication of a strong relationship between "seat-time" and attrition.

- **Aptitude**- Aptitude level was found to be inversely correlated with attrition. However, this relationship was weaker among the non-high school graduates. Additionally, the relationship between aptitude and attrition weakened in the more recent cohorts. This was especially evident in the attrition behavior of recruits that survived beyond boot camp.

- **Moral Waivers**- The presence of a moral waiver corresponded to a slightly higher attriticn rate. This relationship was consistent for all education groups.

B. ATTRITION MODEL

The survivor functions are modelled using a log-linear regression model with the Weibull as the underlying distribution. The results of the model compare favorably to the *product-limit estimates* of the survivor functions. Some of the advantages of using the model to analyze the effects of the covariates on attrition are:

- Computational efficiency
- Easily accommodates additional covariates
- Handles non-categorical covariates

Some of the disadvantages of using the model are:

- Loss of precision when the actual survivor function is ill-behaved
- Accuracy of results degrades with smaller sample sizes

Essentially then, use of the model implies a tradeoff between time and accuracy. The greatest utility of the model may be as a desk-top tool to perform informal analyses. The *product-limit estimates* would be preferable when formal results are required.

C. RECOMMENDATIONS FOR FURTHER STUDY

The attrition behavior of alternate high school credential holders warrants further study of additional and more recent data. Analyzing the effect of "seat-time" in school on attrition would also be a meaningful undertaking. The weakening of aptitude as a predictor of attrition should be more closely examined. Perhaps the behavior is the result of some policy changes within the Marine Corps during the period. Or possibly the behavior is an indicator of social change, in which case similar patterns probably exist in the other services.

Cox's regression is a nonparametric technique that handles censored failure-time data [Ref. 11]. The use of Cox's Regression on this dataset would provide a means to evaluate its effectiveness in military attrition studies.

APPENDIX A. COVARIATE DETAILS

A. AFMG

Recruits are categorized into an Armed Forces Mental Group based on their AFQT percentile as follows:

- AFMG I: 93-99
- AFMG II: 65-92
- AFMG IIIA: 50-64
- AFMG IIIB: 31-49
- AFMG IVA: 21-30
- AFMG IVB: 10-20
- AFMG V: 1 [illegible]

B. RECOGNIZED ALTERNATE CREDENTIALS

The following credentials represent those present in the data that are recognized under the DOD TIER 2 classification:

- **High School Certificate of Attendance:** awarded to regular full-time members of a graduating class who have failed to complete specified requirements for graduation (e.g., minimum number of credit units, minimum GPA, competency test).
- **Correspondence School Certificate:** awarded to graduates of home study schools. While correspondence schools share many of the typical high school academic and course requirements, they lack their attendance requirements and resulting social experiences.
- **GED:** General Education Development (GED) high school equivalency credentials are awarded to those who achieve state-set minimum passing scores on a battery of five tests developed by the American Council on Education. The GED tests are normed on high school seniors and these norms are periodically updated.
- **Occupational Program Certificate:** awarded for attending a non-correspondence vocational, technical, or proprietary school for at least six months, following eleven years or more of traditional schooling.

C. MORAL WAIVERS

Moral waivers recognized in this study are those awarded at the DEPOT or CMC level for any of the following reasons:

- Minor traffic offenses
- Minor non-traffic offenses
- Other (nonminor) offenses
- Felony (adult)
- Felony (juvenile)
- Preservice illegal use of drugs
- Preservice alcohol abuse
- Other (misc.)

APPENDIX B. ADDITIONAL FIGURES

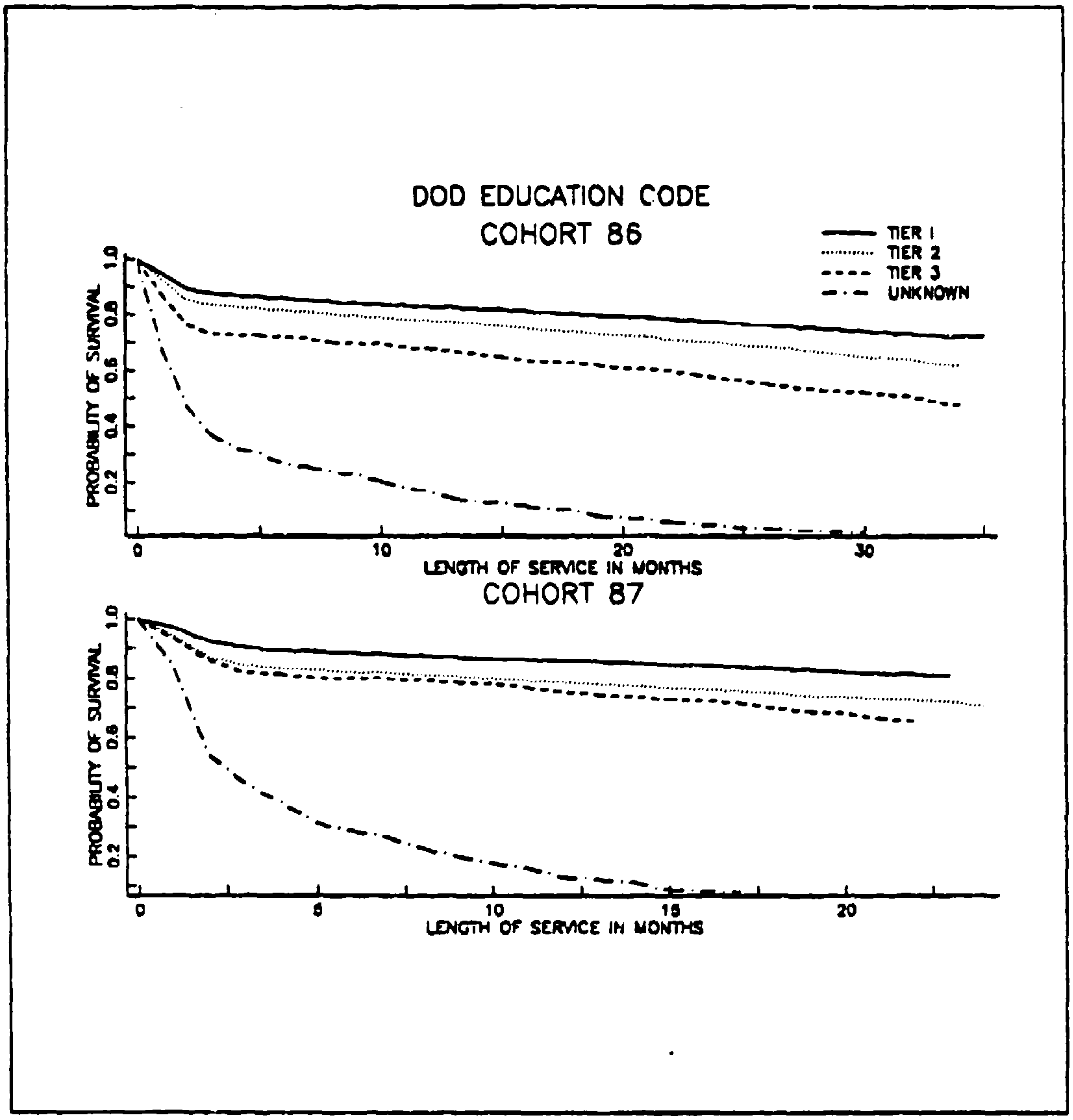

Figure 24. Survivor Functions by DOD Education Code, Cohorts 86 and 87

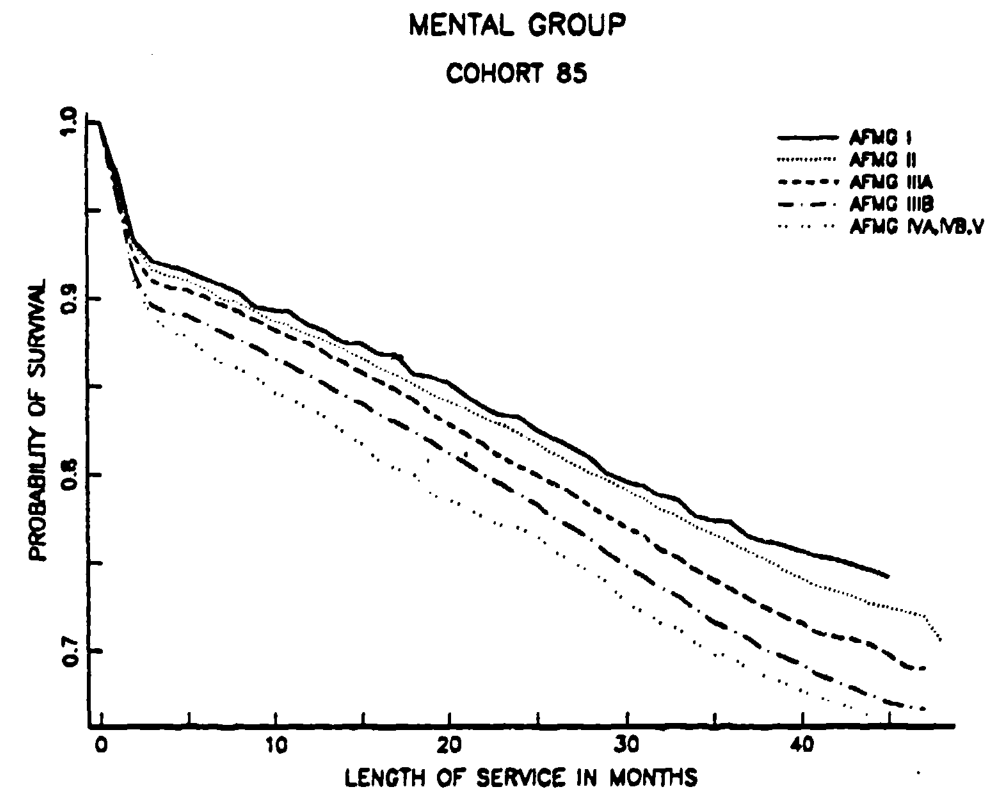

Figure 25. Survivor Functions by Mental Group, Cohort 85

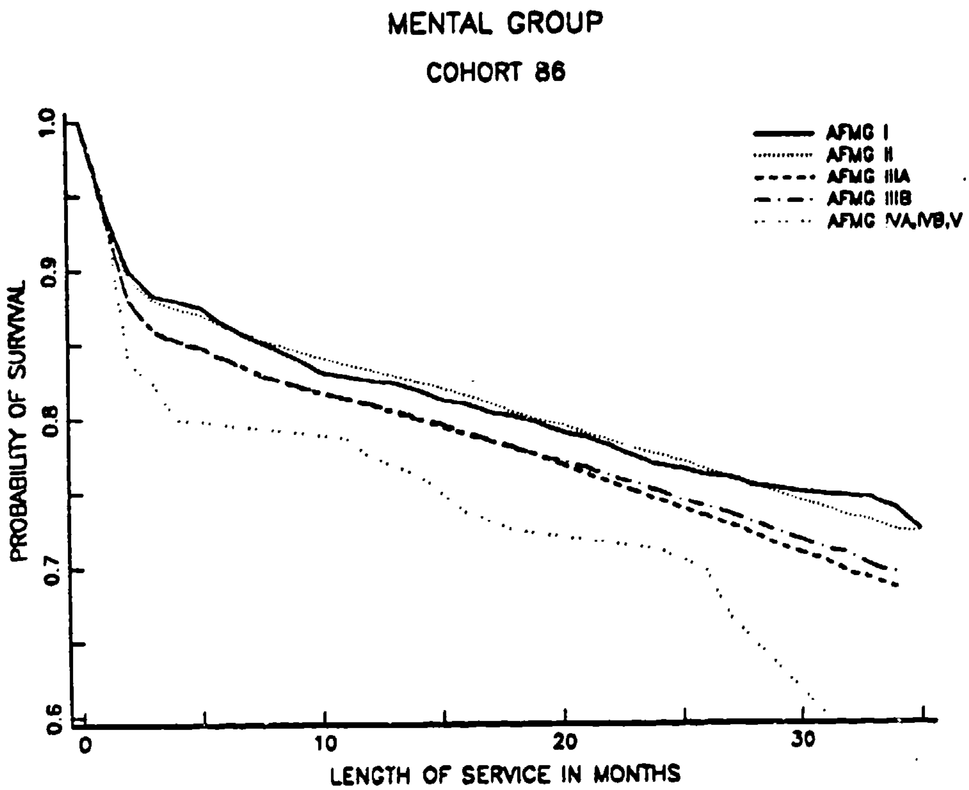

Figure 26. Survivor Functions by Mental Group, Cohort 86

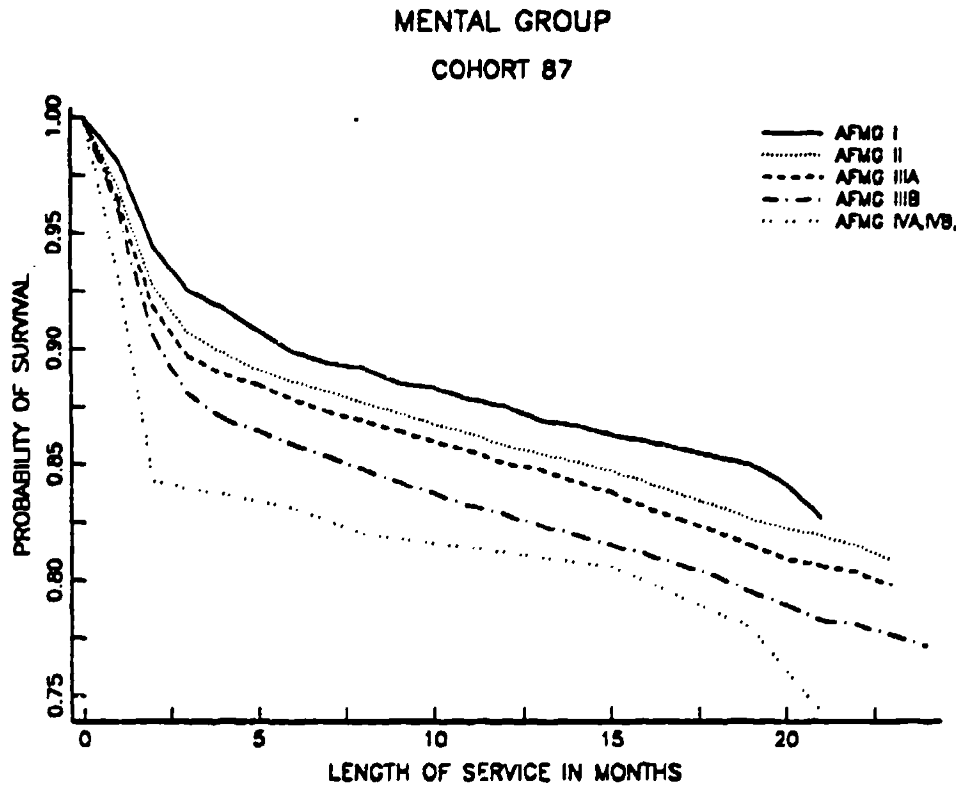

Figure 27. Survivor Functions by Mental Group, Cohort 87

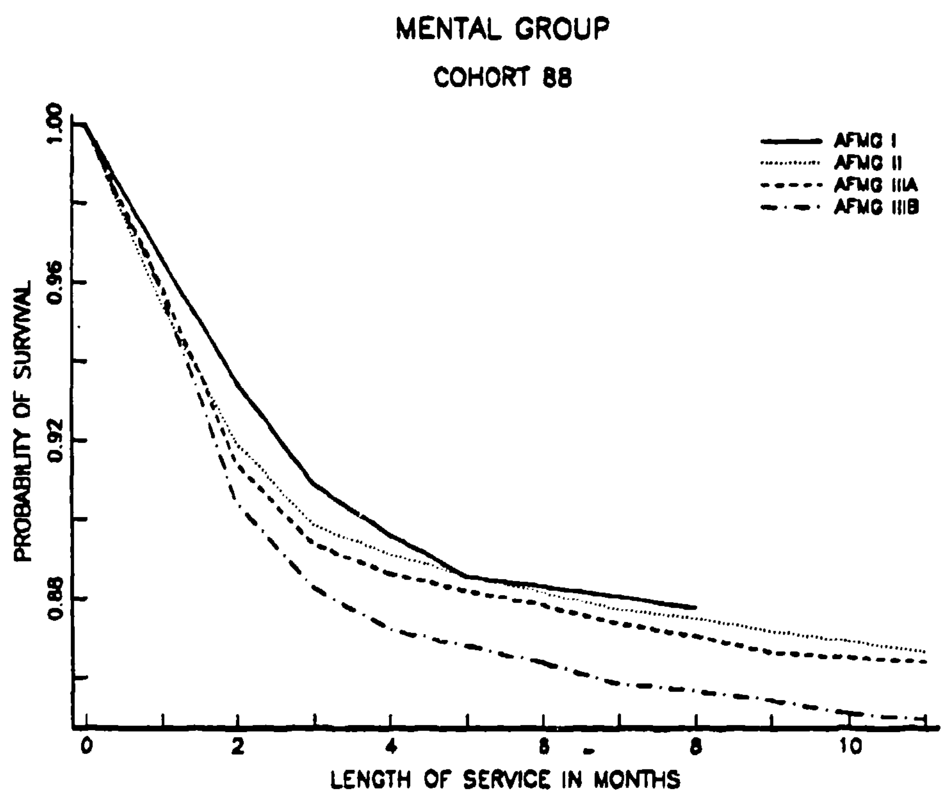

Figure 28. Survivor Functions by Mental Group, Cohort 88

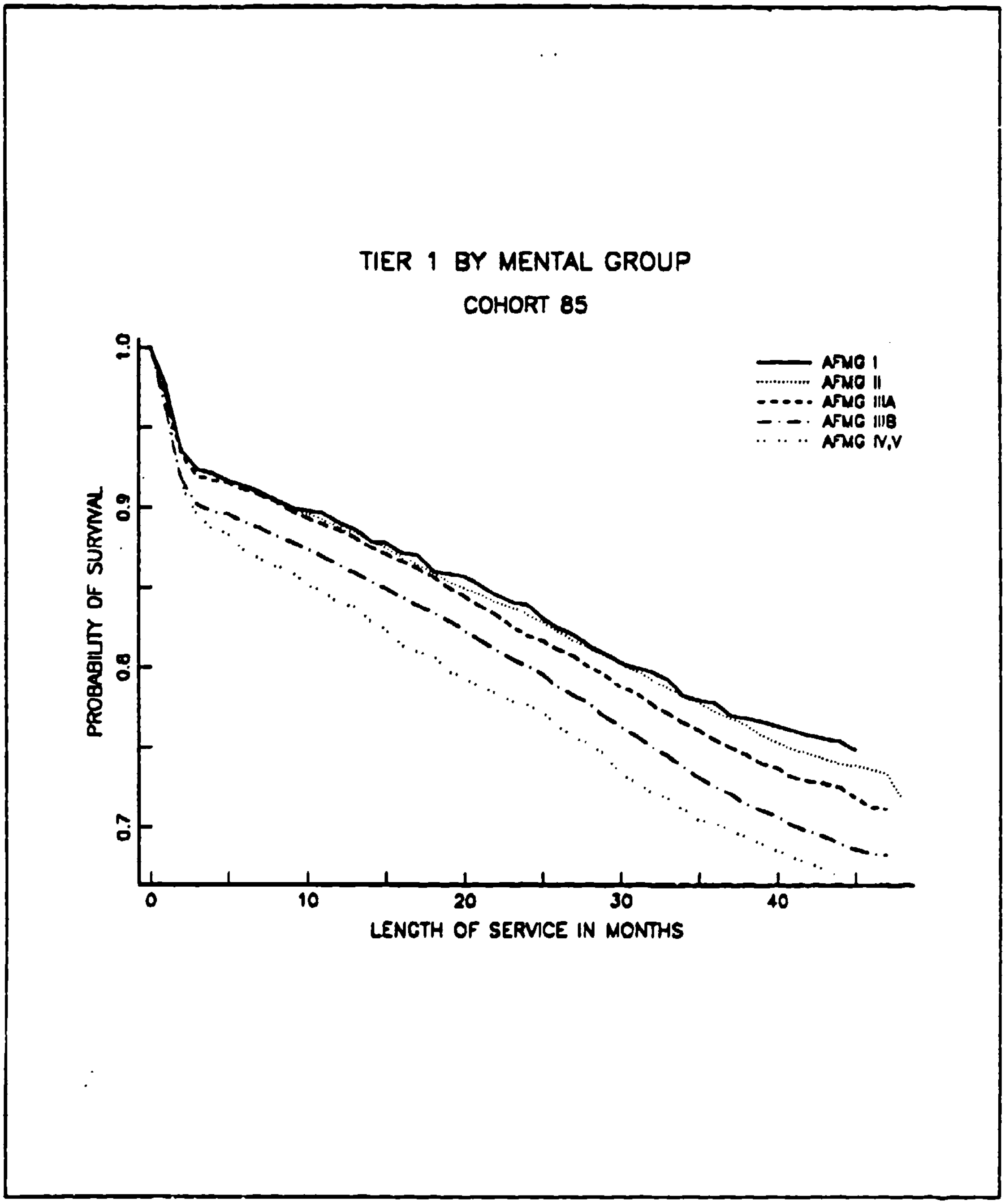

Figure 29. Survivor Functions, TIER 1 by Mental Group, Cohort 85

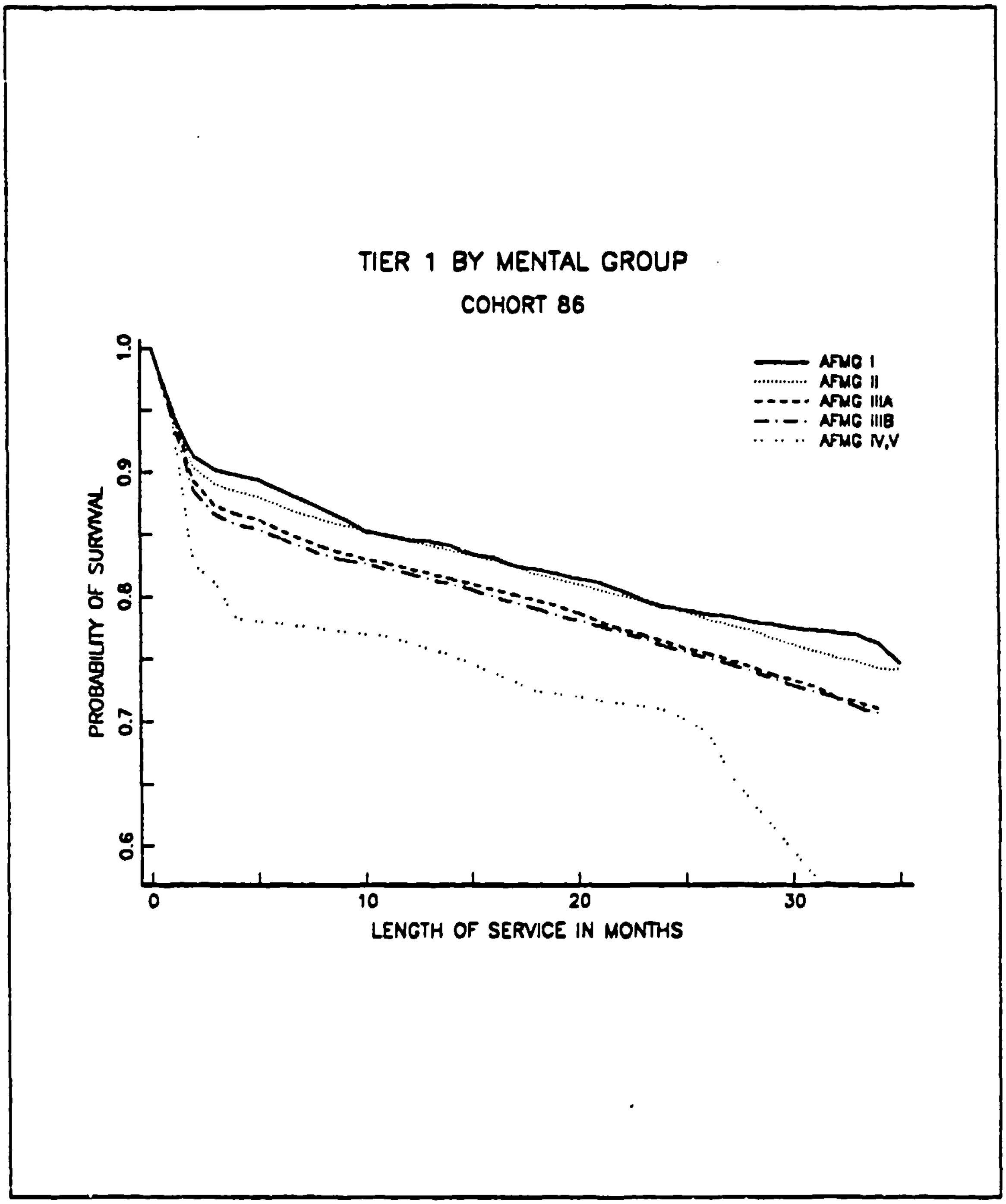

Figure 30. Survivor Functions, TIER 1 by Mental Group, Cohort 86

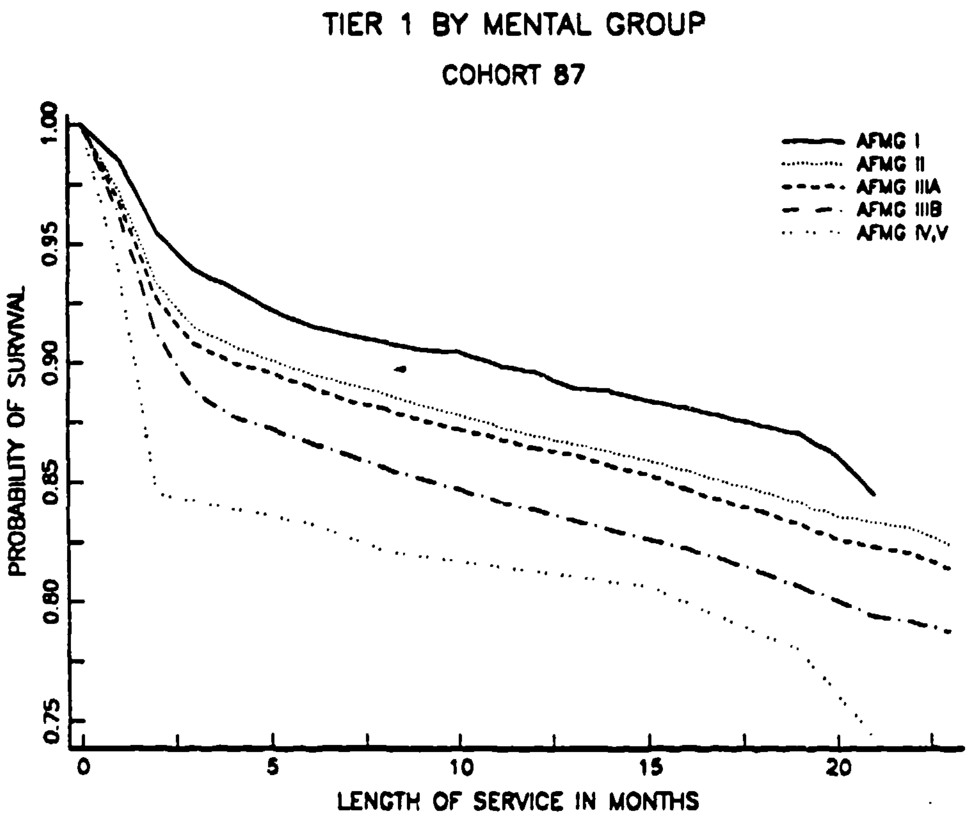

Figure 31. Su"vivor Functions, TIER 1 by Mental Group, Cohort 87

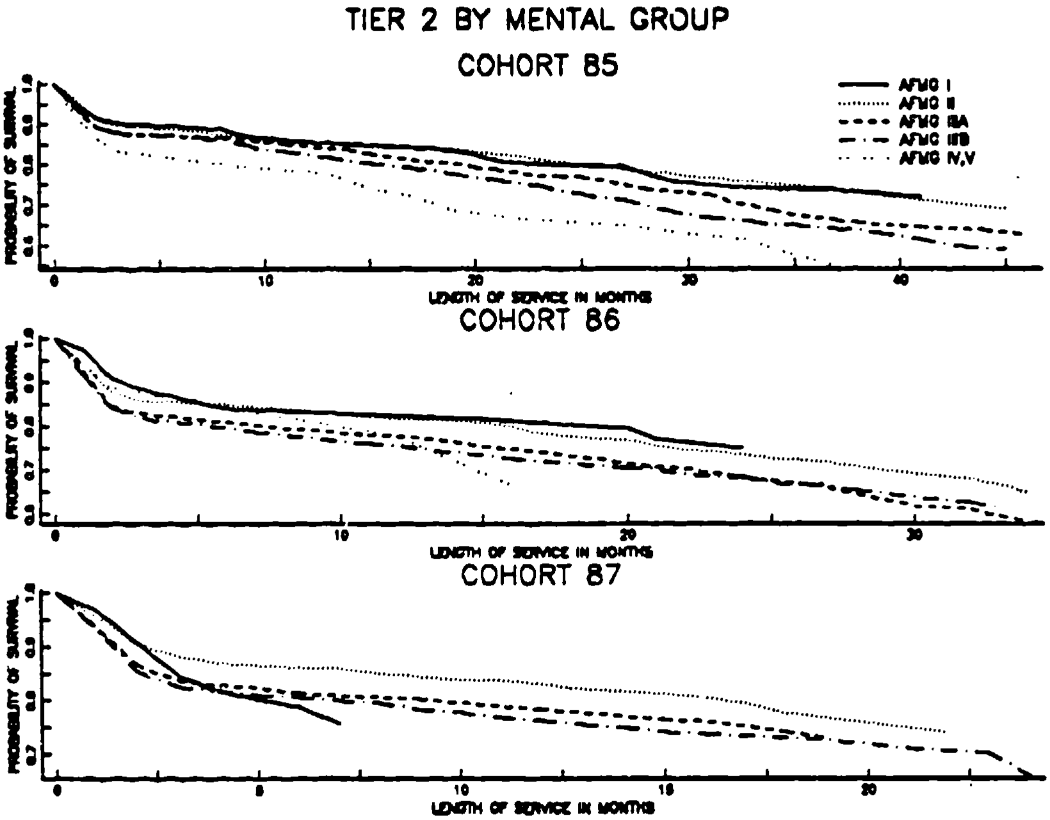

Figure 32. Survivor Functions, TIER 2 by Mental Group, Cohorts 86, 87 and 88

APPENDIX C. SAS OUTPUT FOR PRODUCT-LIMIT ESTIMATES

PRODUCT LIMIT SURVIVAL ESTIMATES
COHORT=84

MONTH	SURVIVAL	FAILURE	STANDARD ERROR	NUM. EVENT	NUM. LEFT
0.0000	1.0000	0.0000	0.0000	0	36079
1.0000	0.9494	0.0506	0.0012	1826	34253
2.0000	0.8938	0.1062	0.0016	3830	32249
3.0000	0.8713	0.1287	0.0018	4645	31434
4.0000	0.8660	0.1340	0.0018	4836	31243
5.0000	0.8640	0.1360	0.0018	4907	31172
6.0000	0.8616	0.1384	0.0018	4995	31083
7.0000	0.8584	0.1416	0.0018	5108	30966
8.0000	0.8555	0.1445	0.0019	5214	30859
9.0000	0.8523	0.1477	0.0019	5328	30742
10.0000	0.8487	0.1513	0.0019	5457	30613
11.0000	0.8451	0.1549	0.0019	5587	30482
12.0000	0.8414	0.1586	0.0019	5722	30346
13.0000	0.8371	0.1629	0.0019	5876	30192
14.0000	0.8332	0.1668	0.0020	6019	30048
15.0000	0.8296	0.1704	0.0020	6148	29919
16.0000	0.8251	0.1749	0.0020	6308	29759
17.0000	0.8213	0.1787	0.0020	6447	29620
18.0000	0.8167	0.1833	0.0020	6612	29455
19.0000	0.8118	0.1882	0.0021	6790	29276
20.0000	0.8065	0.1935	0.0021	6979	29083
21.0000	0.8012	0.1988	0.0021	7170	28892
22.0000	0.7955	0.2045	0.0021	7376	28686
23.0000	0.7907	0.2093	0.0021	7549	28512
24.0000	0.7852	0.2148	0.0022	7748	28312
25.0000	0.7798	0.2202	0.0022	7945	28114
26.0000	0.7745	0.2255	0.0022	8134	27924
27.0000	0.7687	0.2313	0.0022	8344	27713
28.0000	0.7636	0.2364	0.0022	8529	27525
29.0000	0.7586	0.2414	0.0023	8707	27346
30.0000	0.7533	0.2467	0.0023	8898	27154
31.0000	0.7486	0.2514	0.0023	9068	26983
32.0000	0.7432	0.2568	0.0023	9261	26790
33.0000	0.7380	0.2620	0.0023	9450	26601
34.0000	0.7328	0.2672	0.0023	9639	26408
35.0000	0.7282	0.2718	0.0023	9804	26169
36.0000	0.7232	0.2768	0.0024	9980	25830
37.0000	0.7185	0.2815	0.0024	10146	24921
38.0000	0.7143	0.2857	0.0024	10290	24627
39.0000	0.7097	0.2903	0.0024	10448	24439
40.0000	0.7054	0.2946	0.0024	10594	24274
41.0000	0.7018	0.2982	0.0024	10720	24140
42.0000	0.6976	0.3024	0.0024	10864	23988
43.0000	0.6943	0.3057	0.0024	10978	23866
44.0000	0.6903	0.3097	0.0024	11115	23724
45.0000	0.6866	0.3134	0.0025	11243	23591
46.0000	0.6823	0.3177	0.0025	11387	23284
47.0000	0.6786	0.3214	0.0025	11503	21240
48.0000	0.6701	0.3299	0.0025	11745	18961
49.0000	0.6654	0.3346	0.0025	11814	9752
59.0000*	.	.	.	11990	0

* CENSORED OBSERVATION

PRODUCT LIMIT SURVIVAL ESTIMATES
COHORT=85

MONTH	SURVIVAL	FAILURE	STANDARD ERROR	NUM. EVENT	NUM. LEFT
0.0000	1.0000	0.0000	0.0000	0	28944
1.0000	0.9613	0.0387	0.0011	1119	27825
2.0000	0.9211	0.0789	0.0016	2283	26661
3.0000	0.9051	0.0949	0.0017	2746	26196
4.0000	0.9013	0.0987	0.0018	2858	26081
5.0000	0.8992	0.1008	0.0018	2917	26021
6.0000	0.8950	0.1050	0.0018	3040	25898
7.0000	0.8902	0.1098	0.0018	3179	25755
8.0000	0.8858	0.1142	0.0019	3306	25625
9.0000	0.8809	0.1191	0.0019	3446	25484
10.0000	0.8759	0.1241	0.0019	3591	25339
11.0000	0.8716	0.1284	0.0020	3717	25211
12.0000	0.8666	0.1334	0.0020	3860	25065
13.0000	0.8618	0.1382	0.0020	3998	24926
14.0000	0.8567	0.1433	0.0021	4148	24776
15.0000	0.8514	0.1486	0.0021	4301	24622
16.0000	0.8462	0.1538	0.0021	4450	24472
17.0000	0.8413	0.1587	0.0021	4591	24329
18.0000	0.8360	0.1640	0.0022	4745	24173
19.0000	0.8302	0.1698	0.0022	4913	24004
20.0000	0.8245	0.1755	0.0022	5077	23840
21.0000	0.8189	0.1811	0.0023	5239	23678
22.0000	0.8133	0.1867	0.0023	5401	23515
23.0000	0.8078	0.1922	0.0023	5560	23355
24.0000	0.8027	0.1973	0.0023	5708	23206
25.0000	0.7972	0.2028	0.0024	5668	23044
26.0000	0.7905	0.2095	0.0024	6062	22849
27.0000	0.7849	0.2151	0.0024	6224	22686
28.0000	0.7789	0.2211	0.0024	6395	22514
29.0000	0.7726	0.2274	0.0025	6579	[illegible]
30.0000	0.7657	0.2343	0.0025	6778	[illegible]
31.0000	0.7604	0.2396	0.0025	6931	[illegible]
32.0000	0.7541	0.2459	0.0025	7114	[illegible]
33.0000	0.7486	0.2514	0.0026	7271	21636
34.0000	0.7423	0.2577	0.0026	7453	21454
35.0000	0.7368	0.2632	0.0026	7612	21271
36.0000	0.7320	0.2680	0.0026	7752	21081
37.0000	0.7267	0.2733	0.0026	7888	18749
38.0000	0.7218	0.2782	0.0026	8000	16381
39.0000	0.7164	0.2836	0.0027	8105	13979
40.0000	0.7126	0.2874	0.0027	8168	11818
41.0000	0.7071	0.2929	0.0028	8252	10875
42.0000	0.7037	0.2963	0.0028	8299	9684
43.0000	0.7009	0.2991	0.0028	8333	8475
44.0000	0.6970	0.3030	0.0029	8373	7090
45.0000	0.6932	0.3068	0.0029	8402	5364
46.0000	0.6903	0.3097	0.0030	8419	3986
47.0000	0.6886	0.3114	0.0031	8424	2100
48.0000	0.6844	0.3156	0.0052	8425	163
58.0000*	.	.	.	8425	0

* CENSORED OBSERVATION

PRODUCT LIMIT SURVIVAL ESTIMATES
COHORT=86

MONTH	SURVIVAL	FAILURE	STANDARD ERROR	NUM. EVENT	NUM. LEFT
0.0000	1.0000	0.0000	0.0000	0	30346
1.0000	0.9430	0.0570	0.0013	1729	28617
2.0000	0.8853	0.1147	0.0018	3480	26866
3.0000	0.8675	0.1325	0.0019	4022	26324
4.0000	0.8606	0.1394	0.0020	4229	26116
5.0000	0.8561	0.1439	0.0020	4366	25977
6.0000	0.8486	0.1514	0.0021	4594	25748
7.0000	0.8417	0.1583	0.0021	4804	25534
8.0000	0.8355	0.1645	0.0021	4993	25345
9.0000	0.8307	0.1693	0.0022	5136	25202
10.0000	0.8261	0.1739	0.0022	5277	25060
11.0000	0.8218	0.1782	0.0022	5408	24929
12.0000	0.8180	0.1820	0.0022	5524	24812
13.0000	0.8136	0.1864	0.0022	5655	24680
14.0000	0.8090	0.1910	0.0023	5797	24537
15.0000	0.8043	0.1957	0.0023	5937	24395
16.0000	0.7992	0.2008	0.0023	6093	24238
17.0000	0.7944	0.2056	0.0023	6237	24094
18.0000	0.7891	0.2109	0.0023	6400	23931
19.0000	0.7838	0.2162	0.0024	6559	23772
20.0000	0.7792	0.2208	0.0024	6698	23632
21.0000	0.7743	0.2257	0.0024	6848	23482
22.0000	0.7687	0.2313	0.0024	7017	23312
23.0000	0.7632	0.2368	0.0024	7185	23143
24.0000	0.7581	0.2419	0.0025	7338	22986
25.0000	0.7528	0.2472	0.0025	7484	20573
26.0000	0.7478	0.2522	0.0025	7603	17834
27.0000	0.7424	0.2576	0.0025	7714	15105
28.0000	0.7373	0.2627	0.0026	7800	12609
29.0000	0.7309	0.2691	0.0026	7895	10866
30.0000	0.7256	0.2744	0.0027	7963	9205
31.0000	0.7205	0.2795	0.0028	8017	7609
32.0000	0.7146	0.2854	0.0029	8066	6007
33.0000	0.7094	0.2906	0.0030	8103	4973
34.0000	0.7035	0.2965	0.0031	8132	3490
35.0000	0.7028	0.2972	0.0032	8134	1894
43.0000*	.	.	.	8134	0

* CENSORED OBSERVATION

PRODUCT LIMIT SURVIVAL ESTIMATES
COHORT=87

MONTH	SURVIVAL	FAILURE	STANDARD ERROR	NUM. EVENT	NUM. LEFT
0.0000	1.0000	0.0000	0.0000	0	31626
1.0000	0.9655	0.0345	0.0010	1092	30534
2.0000	0.9177	0.0823	0.0015	2603	29023
3.0000	0.8958	0.1042	0.0017	3295	28331
4.0000	0.8866	0.1134	0.0018	3587	28039
5.0000	0.8805	0.1195	0.0018	3778	27846
6.0000	0.8749	0.1251	0.0019	3955	27663
7.0000	0.8699	0.1301	0.0019	4114	27503
8.0000	0.8653	0.1347	0.0019	4260	27355
9.0000	0.8602	0.1398	0.0020	4421	27193
10.0000	0.8557	0.1443	0.0020	4562	27050
11.0000	0.8510	0.1490	0.0020	4712	26899
12.0000	0.8466	0.1534	0.0020	4852	26758
13.0000	0.8426	0.1574	0.0021	4967	24268
14.0000	0.8385	0.1615	0.0021	5070	21000
15.0000	0.8342	0.1658	0.0021	5160	17642
16.0000	0.8295	0.1705	0.0022	5245	14961
17.0000	0.8245	0.1755	0.0022	5326	13497
18.0000	0.8199	0.1801	0.0023	5393	11899
19.0000	0.8137	0.1863	0.0024	5470	10122
20.0000	0.8083	0.1917	0.0025	5525	8240
21.0000	0.8035	0.1965	0.0026	5563	6288
22.0000	0.8006	0.1994	0.0027	5580	4748
23.0000	0.7955	0.2045	0.0029	5597	2641
24.0000	0.7938	0.2062	0.0034	5598	466
26.0000	0.7577	0.2423	0.0354	5599	21
58.0000*	.	.	.	5599	0

* CENSORED OBSERVATION

PRODUCT LIMIT SURVIVAL ESTIMATES
COHORT=88

MONTH	SURVIVAL	FAILURE	STANDARD ERROR	NUM. EVENT	NUM. LEFT
0.0000	1.0000	0.0000	0.0000	0	14319
1.0000	0.9576	0.0424	0.0017	607	13712
2.0000	0.9131	0.0869	0.0024	1245	13074
3.0000	0.8920	0.1080	0.0026	1546	12773
4.0000	0.8835	0.1165	0.0027	1668	12651
5.0000	0.8786	0.1214	0.0027	1739	12580
6.0000	0.8746	0.1254	0.0028	1796	12520
7.0000	0.8699	0.1301	0.0028	1853	10525
8.0000	0.8677	0.1323	0.0029	1875	8704
9.0000	0.8645	0.1355	0.0029	1900	6817
10.0000	0.8621	0.1379	0.0030	1914	4989
11.0000	0.8602	0.1398	0.0031	1920	2807
44.0000*	.	.	.	1920	0

* CENSORED OBSERVATION

EVENTS	CENSORED	TOTAL	%CENSORED		STRATA
11990	24089	36079	66.7674		84
8425	20519	28944	70.8921		85
8134	22212	30346	73.1958		86
5599	26027	31626	82.2962		87
1920	12399	14319	86.5912		88
36068	105246	141314	74.4767	TOTAL	

APPENDIX D. SAS OUTPUT FOR REGRESSION MODEL

L I F E R E G   P R O C E D U R E

CLASS LEVEL INFORMATION

CLASS	LEVELS	VALUES
ED_NEW	4	11 22 33 44
SAS		2

L I F E R E G   P R O C E D U R E

BOOT CAMP 86

```
DATA SET           =WORK.OMEGA
DEPENDENT VARIABLE=EXADAYS
CENSORING VARIABLE=CENSOR
CENSORING VALUE(S)=      1
NONCENSORED VALUES=4229  CENSORED VALUES=    3

LOGLIKELIHOOD FOR WEIBULL        -3888.56
```

VARIABLE	DF	ESTIMATE	STD ERR	CHISQUARE	PR>CHI	LABEL/VALUE
INTERCPT	1	3.89401	.0568874	4685.55	0.0001	INTERCEPT
ED_NEW	3			13.5612	0.0036	
	1	-0.0247283	.0480343	0.265024	0.6067	11
	1	-0.0567739	.0543521	1.0911	0.2962	22
	1	-0.171023	.0620429	7.59841	0.0058	33
	0	0	0	.	.	44
SCORE	1	-.00058187	4.8E-04	1.46322	0.2264	
WAIVER	1	-0.108065	0.030634	12.4441	0.0004	
SCALE	1	0.542358	.0062173			EXTREME VALUE SCALE PARAMETER

L I F E R E G   P R O C E D U R E

COHORT 86

DATA SET =WORK.OMEGA
DEPENDENT VARIABLE=XMON
CENSORING VARIABLE=CENSOR
CENSORING VALUE(S)= 1
NONCENSORED VALUES= 3905 CENSORED VALUES= 22209

LOGLIKELIHOOD FOR WEIBULL -14408.4

VARIABLE	DF	ESTIMATE	STD ERR	CHISQUARE	PR>CHI	LABEL/VALUE
INTERCPT	1	2.38534	0.128727	343.368	0.0001	INTERCEPT
ED_NEW	3		.	473.48	0.0001	
	1	2.38683	0.12106	388.723	0.0001	11
	1	2.05947	0.126968	263.103	0.0001	22
	1	1.70447	0.1403	147.593	0.0001	33
	0	0	0	.	.	44
SCORE	1	0.00231455	8.1E-04	8.25658	0.0041	
WAIVER	1	-0.271454	.0486777	31.0981	0.0001	
SCALE	1	0.883645	.0136557			EXTREME VALUE SCALE PARAMETER

LIST OF REFERENCES

1. Human Resources Research Organization Report 84-4, *Education Standards for Military Enlistment and the Search for Successful Recruits*, by J.H. Laurence, February 1984.

2. Defense Personnel Security Research and Education Center Report PERS-TR-88-006, *Moral Waivers as Predictors of Unsuitability Attrition in the Military*, by C.C. Fitz and M.A. McDaniel, December 1988.

3. Human Resources Research Organization Report 84-6, *Characteristics and Performance of Recruits With General Education Development (GED) Credentials*, by B. Means and J.H. Laurence, May 1984.

4. Office of Naval Research (TR-3), *A Longitudinal Study of Enlisted Personnel Attrition in the U.S. Marine Corps: Preliminary Recruit Training Results*, by W.T. Matthews, June 1977.

5. Office of Naval Research (TR-10), *Involuntary Attrition: Predictions for Non-High-School-Graduate Army Enlistees*, by J.S. Blandin and J.H. Morris, October 1981.

6. Kaplan, E.L., and Meier, P., "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, v. 53, pp. 457-481, 1958.

7. Kalbfleisch, J.D., and Prentice, R.L., *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, Inc., 1980.

8. Gillespie, M.J., and Fisher, L., "Confidence Bands for the Kaplan-Meier Survival Curve Estimate," *Annals of Statistics*, v. 7, pp. 920-924, 1979.

9. Bartholomew, D.J., and Forbes, A.F., *Statistical Techniques for Manpower Planning*, John Wiley & Sons, Inc., 1979.

10. SAS Institute Inc., *SAS User's Guide: Statistics*, Version 5 Edition, 1985.

11. Center for Naval Analyses CRC 402, *Nonparametric Methods for Estimating Recruit Survival with Cross-Sectional Data*, by P.M. Lurie, September 1979.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	Commandant of the Marine Corps Code TE 06 Headquarters, U.S. Marine Corps Washington, D.C. 20380-0001	2
4.	Major J.P. Curry Headquarters, U.S. Marine Corps (Code MA) Washington, D.C. 20380-0001	4
5.	Professor R.R. Read (Code 55RE) Operations Analysis Dept. Naval Postgraduate School Monterey, CA 93943	2
6.	Professor P.R. Milch (Code 55MH) Operations Analysis Dept. Naval Postgraduate School Monterey, CA 93943	2

END
FILMED
12-90
DTIC